UNIVERSITY of CALIFORNIA
Santa Barbara

Testing galaxy formation and the nature of dark matter with satellite galaxies

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Physics

by

Anna Mercedes Nierenberg

Committee in charge:

Professor Tommaso Treu, Chair
Professor Crystal Martin
Professor Peng Oh

September 2014

The dissertation of Anna Mercedes Nierenberg is approved:

Professor Crystal Martin

Professor Peng Oh

Professor Tommaso Treu, Chair

September 2014

Testing galaxy formation and the nature of dark matter with satellite galaxies

Anna Mercedes Nierenberg

To my parents; it really is a wonderful world.

Acknowledgments

It has been a long road getting here, and there are a lot of people to whom I owe this achievement. To my parents, thank you for your selfless and limitless love, support and encouragement. You taught me to never settle, to honor hard work in myself and others and to take time to enjoy life. You are both always on my team no matter what, and I am so grateful for that.

To my brother Will, you are such a good friend. Growing up with you has made me a better person. I love and admire you so much and am so grateful to have you in my life. Thank you for helping me learn to ask for what I want, and to be less afraid! Also, thank you for sharing my passionate dislike for orange chocolate, which validates how I feel.

To my grandparents: Nanny I will always treasure how you nurtured my artistic side by drawing and painting with me when I was little. I learned so much from you about how to brighten life with colors. Moosie and Bobo I am truly blessed to know you and see how your love touches so many lives and makes the world a more beautiful place. I am grateful for the example of your marriage, how you have grown closer and closer together, over so many years, and for showing me how important it is in life to have a strong and loving family. And to Grandpa Bill, I am so honored by your example. You couldn't make it to my thesis defense, but I know you've been with me every step of the way.

I am grateful to the many teachers in my life. In particular, thank you to Professor Creswell, who taught my high school calculus class. You are a dedicated and gifted

teacher who made calculus easy, approachable and interesting. Also thank you to Professor Coroniti, my first college physics professor who gave me lots of encouragement.

In graduate school, I am indebted to my cohort, whose moral support and homework collaboration made first year possible, physically and psychologically. I am also grateful to my fellow astronomy graduate students, I will miss you guys, especially at lunch. And to the other two members of the 'A-Team', I look forward to a future of partnership and collaboration! Don't take my head off the map please... at least for a little while.

Thank you to the more senior students and post-docs, who have supported and advised me. I knew I could count on you any time and I am glad I had such positive role models to light the road ahead.

Of course, none of this would have been possible without my wonderful advisor, Tommaso. I literally couldn't have imagined a better advisor, and I am a pretty creative person. Thank you for your endless patience, encouragement, and for all the opportunities I've had in my graduate career. You have taught me not only about science but also how to be a scientist. I am still a long way from achieving zen mastery, but thanks to you I have seen how it's possible for a person to care a lot about something but still react calmly in the face of the randomness of life. As an unexpected bonus, in the six years of finishing this dissertation, I have learned WAY more about food than I ever expected to. I look forward to much continued collaboration, and many delicious meals in the coming years.

Finally, to Daniel, I am so glad I got to share this experience with you. It has been a long six years and you've been with me every step of the way, with hugs when I'm down and dancing when I'm up!

Curriculum Vitae

Anna Mercedes Nierenberg

Education

2014	Ph.D. in Physics, University of California, Santa Barbara
2008	B.S. in Physics, University of California, Los Angeles, *summa cum laude*

Professional Experience

2009-2014	Graduate Research Assistant, Department of Physics, University of California, Santa Barbara
2008-2009	Teaching Associate, University of California, Santa Barbara
2007-2008	Undergraduate student researcher, University of California, Los Angeles
Summer 2007	Undergraduate student researcher, Scripps Institution of Oceanography

First Author Publications

"Detection of substructure with adaptive optics integral field spectroscopy of the gravitational lens B1422+231", Nierenberg, A. M.; Treu, T.; Wright, S. A.; Fassnacht, C. D.; Auger, M. W.; 2014, MNRAS 43, 2120 (2014)

"Do lens galaxies have an excess of luminous substructure?", Nierenberg, A. M.; Oldenburg, D.; Treu, T.; MNRAS, 436, 2120 (2013)

"The Cosmic Evolution of Faint Satellite Galaxies as a Test of Galaxy Formation and the Nature of Dark Matter", Nierenberg, A. M.; Treu, T.; Menci, N.; Lu, Y.; Wang, W.; ApJ, 772, 146, (2013)

"Luminous Satellites of Early-Type Galaxies II: Spatial Distribution, Luminosity Function and Cosmic Evolution", Nierenberg, A. M.; Auger, M. W.; Treu, T.; Marshall, P. J.; Fassnacht, C. D.; Busha, M. T; ApJ 752, 99, (2012)

"Luminous Satellites of Early-Type Galaxies I: Spatial Distribution", Nierenberg, A. M.; Auger, M. W.; Treu, T.; Marshall, P. J.; Fassnacht, C. D.; ApJ 731, 44, (2011)

Fellowships and Awards

2014-2017	CCAPP Postdoctoral Fellow, Ohio State University
2012-2013	Dean's Fellowship, University of California, Santa Barbara
Summer 2012	Worster Undergraduate Research Fellowship, University of California, Santa Barbara
Spring 2012	Physics Chair's Fellowship, University of California, Santa Barbara

Grants

2014	**HST-GO-13732**: Detecting dark matter substructure with narrow line lensing (tbd, PI)
2013	**HST-AR-13271**: The cosmic evolution of faint satellites as a test of galaxy formation and the nature of dark matter ($59,784, co-PI)

Abstract

Testing galaxy formation and the nature of dark matter with satellite galaxies

by

Anna Mercedes Nierenberg

The abundance of low mass halos is one of the key predictions of Λ CDM, which remains at apparent odds with observations of luminous structure. We present new measurements of the spatial distribution and the cumulative luminosity function of satellite galaxies up to a thousand times fainter than their hosts, as a function of host stellar mass and morphology between redshifts 0.1 and 0.8, using imaging from the COSMOS and GOODS fields in conjunction with a rigorous statistical analysis. We demonstrate how these measurements provide powerful new constraints for abundance matching and cosmological simulations in the context of both warm and cold dark matter, and how future measurements of faint satellite colors using CANDELS, will provide important distinguishing power between warm and cold dark matter models. In addition, we present results from a complementary gravitational lens modeling project in which we use strongly lensed AGN narrow-line emission in order to detect dark matter subhalos, demonstrating a promising new method for measuring the subhalo mass function in thousands of lensed systems which will be discovered in ongoing and future optical surveys.

Contents

1 Introduction: Cold dark matter and satellite galaxies 1
1.1 Introduction to luminous satellite galaxies . . . 4
1.1.1 The luminosity function of satellite galaxies . . . 5
1.1.2 The spatial distribution of satellite galaxies . . . 7
1.2 Gravitational lensing . . . 11

2 Luminous satellites I: Determining the number and spatial distribution in a Bayesian framework 18
2.1 Host Galaxy Sample . . . 19
2.2 Detection and photometry of close neighbors . . . 22
2.2.1 Host Galaxy B-Spline Models . . . 23
2.2.2 Object Detection and Photometry . . . 24
2.2.3 Objects Detected in Cutout Regions . . . 26
2.3 First Look . . . 30
2.3.1 Radial Distribution . . . 31
2.3.2 Angular Distribution . . . 32
2.4 Joint Modeling of Satellite and Background Galaxy Populations . . . 33
2.4.1 Satellites . . . 34
Satellite Spatial Distribution . . . 34
Spatial Normalization . . . 37
Number of satellites per host . . . 38
2.4.2 Background/Foreground Objects . . . 39
2.4.3 Analysis . . . 41
2.5 Results . . . 41
2.6 Discussion . . . 45
2.7 Summary . . . 53
2.8 Appendix . . . 54
2.8.1 Host Galaxy Light Subtraction . . . 54
Masking . . . 55
B-spline model subtraction . . . 56
2.8.2 Source Extractor Parameters and Photometry Comparison . . . 56
2.8.3 Completeness . . . 58
2.9 Inference methodology . . . 60

3 Luminous satellites II: Cosmic evolution and dependence on host properties of the satellite luminosity function and spatial distribution 63
3.1 Imaging and Catalogs . . . 65
3.2 Host Galaxy Selection . . . 66
3.3 Detection and photometry of close neighbors . . . 68
3.3.1 Object detection and photometry . . . 69
3.3.2 Properties of Objects Detected in Cutout Regions . . . 70
3.4 First Look . . . 71

3.4.1 Distance Scaling and Radial Distribution . . . 72
3.4.2 Angular Distribution . . . 74
3.5 Joint Modeling of Satellite and Background Galaxy Populations . . . 75
3.5.1 Background/Foreground Objects . . . 75
3.6 Theoretical Prediction for the Number of Satellites per Host . . . 77
3.7 Analysis . . . 78
3.8 Results . . . 80
3.8.1 Radial Distribution . . . 80
3.8.2 Angular Distribution . . . 82
3.8.3 Cumulative Luminosity Function . . . 84
3.9 Comparison with previous work . . . 87
3.10 Discussion . . . 92
3.11 Summary . . . 95
3.12 Appendix . . . 98
3.12.1 Source Extractor Parameters and Photometry Comparison . . . 98
3.13 Completeness . . . 101
3.13.1 Full Posterior PDFs . . . 102

4 Luminous satellites III: Comparison of the luminosity function with theoretical predictions 105
4.1 Theoretical Models . . . 106
4.1.1 Cold Dark Matter Models . . . 107
Menci et al. 2012 . . . 107
Guo et al. 2011 . . . 109
Lu et al. 2012 . . . 110
4.1.2 Menci Warm Dark Matter Model . . . 112
4.2 Results . . . 113
4.3 Comparison of simulation properties . . . 115
4.3.1 Unstripped subhalo mass function . . . 116
4.3.2 The halo mass to stellar mass relationship . . . 118
4.3.3 Satellite galaxy colors . . . 120
4.4 Comparison with literature . . . 122
4.4.1 Satellite luminosity function . . . 124
4.4.2 Satellite Colors . . . 125
4.5 Discussion and conclusion . . . 126

5 The luminous satellites of gravitational lenses 129
5.1 Host Galaxy Selection . . . 132
5.1.1 Lens Host Galaxies . . . 132
5.1.2 Non-Lens Host Galaxies . . . 133
5.2 Companion Detection . . . 135
5.3 Inference of the Number of Satellites Per Host . . . 135
5.3.1 Number of satellites as a function of host stellar mass . . . 136
5.3.2 Uncertainty in the virial radius . . . 136
5.3.3 Priors . . . 137

5.4 Results . . . 138
5.5 Comparison with CLASS . . . 140
5.6 Summary . . . 144
5.7 Appendix: Likelihood function . . . 146

6 Detection of dark matter halos with strong lensing of narrow-line emission 149
6.1 Observations . . . 152
6.2 Data Reduction . . . 152
6.3 Integrated Line Fluxes . . . 154
6.3.1 Astrometry of image D and the main lens galaxy . . . 157
6.4 Gravitational Lens Modelling . . . 158
6.4.1 Singular Isothermal Ellipsoid Deflector . . . 159
6.4.2 Smooth model plus Perturber . . . 161
6.4.3 Finite Source Effects . . . 165
6.5 Results . . . 168
6.6 Discussion . . . 173
6.6.1 Comparison with previous work . . . 175
6.6.2 Limitations of our analysis . . . 176
6.6.3 Future Prospects . . . 177
6.7 Summary . . . 178
6.8 Appendix: Variation in lensing effect with perturber mass profile . . . 180

7 Conclusions and future directions 183
7.1 The colors and stellar mass function of satellite galaxies . . . 184
7.2 Narrow-line lensing going forward . . . 188
7.2.1 Ongoing data analysis . . . 189
7.3 Combining the results . . . 190

Bibliography 192

Chapter 1

Introduction: Cold dark matter and satellite galaxies

The large scale distribution of structure (e.g. Komatsu et al. 2011; Planck Collaboration et al. 2013), the kinematics of galaxy clusters (Clowe et al. 2004), and the rotation curves of spiral galaxies (see e.g. Battaner & Florido 2000, and references therein) can all be explained by assuming that the matter in the Universe is composed mostly of a collisionless particle which does not interact electromagnetically. The success of the dark matter theory is remarkable given that it is simple, while reproducing phenomena which occur across a vast range of distance, mass and time scales. Despite this success, dark matter itself remains elusive, and understanding its detailed physical properties remains one of the most important open problems in modern physics.

There are numerous candidates for a dark matter particle or particles based on extensions to the standard model of particle physics, with the most commonly invoked being weakly interacting massive particle (WIMP). This candidate is appealing because a $\sim$GeV mass particle with a self interaction cross section characteristic of the weak force would naturally produce a dark matter abundance consistent with what is observed (see Strigari 2013, for a more thorough discussion).

One of the key aspects of WIMP dark matter is that its high mass leads to a negligible

free-streaming length so that dark matter structures (halos) form with scale-free power down to extremely small, planet-mass halos. This and other candidates with short free streaming lengths are known as 'cold' dark matter (CDM). CDM is the default model assumed for cosmological simulations and 'coldness' is almost always assumed when interpreting results of direct and indirect detection experiments.

If CDM is the correct model, high resolution cosmological simulations predict that the halo of a galaxy with the mass of the Milky Way should host thousands of self-gravitating dark matter halos embedded in the smooth distribution of the Milky Way dark matter halo. This prediction is in stark contrast with observation, which find only about twenty luminous low mass galaxies gravitationally bound to the Milky Way. This discrepancy was identified originally in the Local Group by Klypin et al. (1999) and Moore et al. (1999), and is observed to extend to the field with the significantly suppressed mass function of low mass field galaxies identified in HI surveys (Papastergis et al. 2011). This discrepancy remains one of the most significant challenges to the CDM theory of dark matter.

There are two non-mutually exclusive classes of solutions for this discrepancy, an astro-particle and an astrophysical solution.

The astrophysical solution attributes the discrepancy to the extremely complex, and not well understood physics of star formation in low mass halos, which may cause a large fraction of subhalos to be too dark to see. While dark matter theory can predict the number of *subhalos*, an understanding of the extremely complex physics of star formation in the shallow potential wells of subhalos is necessary in order to predict how much gas and stars the subhalos should contain. In order to reconcile predicted cold dark matter (CDM)

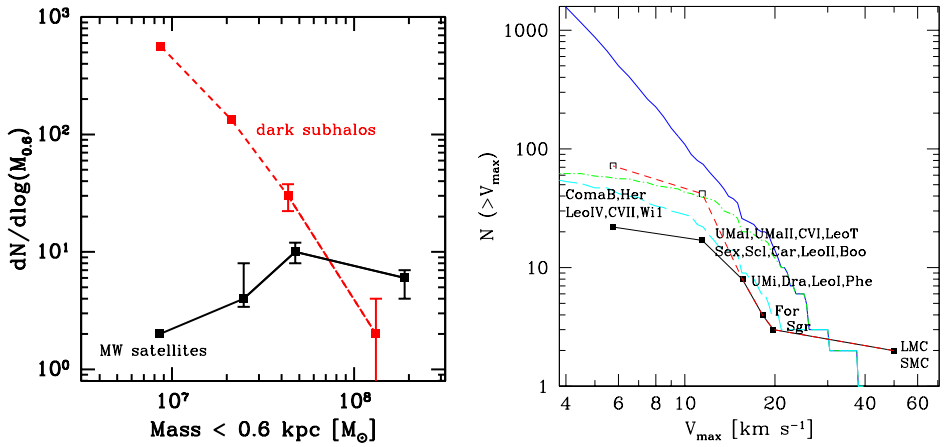

Figure 1.1: Demonstrations of the missing satellite problem. *Left:* Comparison between the differential subhalo mass function for masses measured within 600 pc of the center of luminous Milky Way subhalos, reproduced from Strigari et al. (2007). *Right*: Comparison of predicted (solid blue line) and observed (black points) subhalo cumulative maximum velocity function from Madau et al. (2008). The red dashed line indicates the possible abundance after correcting for the limited sky coverage of the observational sample. There is significant discrepancy between the dark matter model predictions and observation in both cases, even after correcting for sky coverage effects.

halo mass functions at small scales with the observed satellite luminosity function it is necessary to understand the complex interplay of numerous baryonic processes including UV heating during reionization, tidal and ram-pressure stripping by the central galaxy, supernova feedback, and stellar winds (Lu et al. 2012; Menci et al. 2012; Guo et al. 2011a; Bullock et al. 2000; Benson et al. 2002; Somerville 2002; Kravtsov et al. 2004a; Kaufmann et al. 2008; Macciò 2010; Springel 2010). Star formation in satellite galaxies is an interesting problem in its own right with numerous other important astrophysical implications. In Section 1.1 of this Introduction, we discuss in more detail some of the open questions regarding star formation in satellite galaxies and we provide an overview of previous observational constraints.

The astroparticle solutions attribute the observed discrepancy to a fundamental break-

down of the ΛCDM paradigm at small scales, possibly indicating that assumptions about the nature of the CDM particle are incorrect. For instance, DM particles may have higher kinetic energy than expected (e.g. Colín et al. 2000; Schneider et al. 2011; Menci et al. 2012) leading to suppressed fluctuations in the primordial power spectrum at small scales, thereby decreasing the number of present day satellites (e.g. Kamionkowski & Liddle 2000; Zentner & Bullock 2003). The cleanest way to test this latter set of solutions is with a direct measurement of the subhalo mass function which does not rely on an understanding of the baryonic physics in the subhalos. In Section 1.2 of this Introduction, we summarize some of the basic principles of strong gravitational lensing, and discuss how it has been used in the past to detect dark, low mass structure at cosmological distances, potentially enabling a measurement of the subhalo mass function without requiring the subhalos to contain any stars at all.

1.1 Introduction to luminous satellite galaxies

It is important to understand star formation in satellite galaxies not only as a means of understanding CDM, but also because satellite galaxies are believed to have played a crucial role in the size and mass growth of central galaxies, the buildup of stellar halos, and the re-ionization of the Universe (Kaviraj et al. 2011; Bournaud et al. 2007; Conroy et al. 2007; Paardekooper et al. 2013; Bouwens et al. 2012; Kistler et al. 2013). While a variety of theoretical star formation models have been shown to reproduce various properties of observed luminous satellite galaxies, the physics of star formation in satellite galaxies is still not well constrained, with numerous models which use widely different physical

assumptions reproducing the mass function of the Milky Way satellites, for instance (Guo et al. 2011a, see also, Chapter 4).

Baryonic processes are also invoked to resolve the discrepancy between the predicted and observed central velocities of ΛCDM subhalos around the Milky Way (Boylan-Kolchin et al. 2012) by lowering the subhalo central velocities, thereby making them appear less massive. Simulations with baryons have produced conflicting predictions as to whether supernovae can (Governato et al. 2012; Teyssier et al. 2013) or cannot (Garrison-Kimmel et al. 2013) significantly alter the kinematics in the central regions of dark matter halos.

The issue of the low central velocities of Milky Way subhalos is further complicated by the fact that the ΛCDM prediction for the subhalo circular velocity function depends sensitively on the virial mass of the Milky Way halo. Wang et al. (2012) showed that taking into account the uncertainty on the measured Milky Way virial mass, the observed subhalo circular velocity function is consistent with that observed in ensemble simulations of dark matter halos. Purcell & Zentner (2012) further showed that even simulated halos with masses corresponding to the observational mean have subhalo populations consistent with the Milky Way a significant fraction ($\sim 10\%$) of the time.

These studies demonstrate that it is essential to study the statistical properties of satellite galaxies in observations and simulations for a large sample of host galaxies, in order to determine whether apparent discrepancies are driven by the stochastic nature of galaxy formation, or observational uncertainty, rather than poorly modeled physical processes.

Furthermore, as the baryonic processes which affect the subhalo population occur over

cosmological time scales, the predicted effects would ideally be compared with observations of the satellite population over as much of the history of the Universe as possible.

1.1.1 The luminosity function of satellite galaxies

Observations of Milky Way satellites are extremely deep, with satellite galaxies detected which are as faint as $M_V = -2$ (see Koposov et al. 2008, and references therein). Outside of the Local Group, measurements have been restricted to brighter satellites. Multiple studies using data from SDSS have measured the luminosity function of satellites at low redshift (Guo et al. 2011a; Liu et al. 2011; Lares et al. 2011; Strigari & Wechsler 2012). The deepest study by Guo et al. (2011a), measured the luminosity function reliably for satellites up to 7.5 magnitudes fainter than the hosts, while Strigari & Wechsler (2012) placed upper limits on satellite numbers for satellites up to 10 magnitudes fainter. All of these studies agreed on two main conclusions, first that the satellite luminosity function depends on the host luminosity and second that the luminosity function of faint satellites of the Local Group is consistent with that of field galaxies.

The first conclusion, that the satellite luminosity function depends on the host galaxy magnitude, is due to the fact that for massive host galaxies, the host stellar mass is believed to be a non-linear function of the host virial mass (e.g. Behroozi et al. 2010; Wake et al. 2011; Leauthaud et al. 2012). Thus while the number of satellites at a fixed *virial* mass ratio with the host is expected to be scale invariant (Kravtsov 2010), the luminosity ratios may not be. The second conclusion has been tested for satellites as faint as an absolute V band magnitude of -14 (Guo et al. 2011a; Lares et al. 2011;

Strigari & Wechsler 2012), on the bright end only 10-20 percent of hosts with Milky Way and Andromeda luminosities host satellites as luminous as the Large and Small Magellanic Clouds (L/SMCs). The rarity of high mass companions is consistent with predictions from ΛCDM simulations, indicating that the Milky Way luminosity function is an outlier on the bright end rather than that there is an issue with the theory. However, a significant discrepancy between theoretical predictions and observation of the satellite luminosity function only begins to appear for satellites with absolute magnitudes fainter than $M_V > -14$. Thus, in order to test whether the missing satellite problem is due to the Milky Way lacking satellites on the faint end, it is necessary to study even fainter satellites.

In addition to providing a key to understanding star formation in low mass halos, satellite galaxies are believed to have played a significant role in the evolution of the size of the most massive galaxies ($M_* > 10^{10} M_\odot$) (Boylan-Kolchin et al. 2006; Boylan-Kolchin & Ma 2007; Hopkins et al. 2010) [1]. Interestingly recent observations of the evolution of early-type lens galaxies by Sonnenfeld et al. (2014) indicate that the growth of massive ellipticals cannot be explained entirely by dry merging with satellite galaxies, thus a constraint on the gas content, as well as the number of satellite galaxies is important.

The role of minor mergers in the evolution in size and mass has been observed in numerous observational studies which used pair counting and assumptions about merger time-scales to estimate the minor merger rate (Le Fèvre et al. 2000; Bell et al. 2006; Patton & Atfield 2008; Bundy et al. 2009; Bezanson et al. 2009; Robaina et al. 2010; Newman et al. 2012; Tal et al. 2011). Furthermore, measurements of disturbed mor-

[1] But see also Nipoti et al. (2012)

phology and color gradients in an elliptical hosts as evidence for recent mergers (Kaviraj et al. 2009; Kaviraj et al. 2011) show that minor merging is a key contributor to low level star formation seen in the outskirts of early-type galaxies. These studies have all been limited to the study of the most massive companions with stellar masses at least ten percent of their host mass, which is ten times higher than the mass ratio between the Milky Way and the LMC. In order to better constrain the effect of minor mergers on the evolution of massive galaxies, it is necessary to push the study of companions at cosmological distances to lower stellar masses (see also Bundy et al. 2007; Naab et al. 2009; Fakhouri et al. 2010).

1.1.2 The spatial distribution of satellite galaxies

The spatial distribution of satellites about their host galaxies is intimately tied to the luminosity function of those satellite galaxies. Two competing effects contribute to the link between satellite luminosity function and spatial distribution. First, baryons are thought to play an important role in the preservation of subhalos in the inner regions of the host dark matter halo by steepening the total mass profiles and making them less susceptible to tidal stripping. Numerous hydrodynamical and semi-analytic simulations have found that subhaloes with baryons are more centrally located in the host dark matter halo than pure dark matter subhalos (e.g. Blumenthal et al. 1986; Libeskind et al. 2010; Romano-Díaz et al. 2009; Weinberg et al. 2008; Macciò et al. 2006; Gao et al. 2004).

However, baryons will also steepen the mass profile of the host, increasing the strength of tidal shock heating of the gas and stars in satellites as they pass near the center of the

halo, with the amount of heating and mass loss strongly dependent on the orbits of the satellites (Gnedin et al. 1999; Choi et al. 2009; Boylan-Kolchin & Ma 2007; Dolag et al. 2009). D'Onghia et al. (2010) find that the effects of these interactions in suppressing star formation is stronger for less massive subhalos. For these reasons, the slope of the radial profile of satellites is of key importance to any analysis attempting to recover the physics that governs the interactions between satellites and their host halo and central galaxy. Figure 1.2 shows the predicted effect of baryons of satellites around a Milky Way like host galaxy from high resolution hydrodynamic simulations by Brooks & Zolotov (2014).

The radial profile of the number density of satellites has been measured with somewhat contradictory results. Assuming a single power-law model for the projected number density of satellites N as a function of distance from the host, where $N(R) \propto R_p^{\gamma}$, Chen (2008) measured $\gamma_p = -1.5 \pm 0.07$ for satellites brighter than $M_r \sim -17.5$ in SDSS. In a study of SDSS satellites brighter than $M_g < -21.2$, Watson et al. (2010) measured $\gamma_p = -1.2 \pm 0.1$.

There are several likely explanations for the discrepancies in these measurements. One possibility is differences in the technique used to separate satellites from background/foreground interlopers. Alternatively the difference in inferred slopes may be due to differences in host masses studied, or differences in the luminosity of the satellites. For example, Watson et al. (2012) measured the radial profile of satellite galaxies brighter than $M_r < -18$ between 7 and 280 kpc as a function of satellite luminosity, and found that the faintest satellites had significantly shallower radial profiles than their more luminous counterparts. Tal et al. (2012) also found fewer faint satellites in groups

within 25 kpc for satellites brighter than $M_g < -22$. However outside of this region, they found that all satellites followed the same radial profile which was well described by a combined Sersic+NFW profile, mimicking the total mass profile which follows R^{-1} (Gavazzi et al. 2007; Auger et al. 2010) for massive galaxies. In contrast to these studies, Guo et al. (2012) found that for satellites brighter than $m_r < 22$ in SDSS, fainter satellites are *more* centrally concentrated than bright satellites. Budzynski et al. (2012) also found an excess of fainter satellites in the innermost regions for satellites in groups and clusters.

To make progress on understanding the satellite radial profile, it is necessary to make measurements that take into account complexities, such as the host mass and satellite luminosity. In addition, in order to learn about trends in merging rates it is interesting to study evolution in the radial profile. Budzynski et al. (2012) found no significant evolution in the radial profile of satellites in groups and clusters between redshifts 0.15 and 0.4. In this work we extend the measurement of the satellite radial profile to fainter satellites and to higher redshifts.

The angular distribution of satellites also contains information about the effects of anisotropic, filamentary accretion in ΛCDM and the shape of the host dark matter halo. Pure dark matter simulations predict that satellites should be found in an anisotropic distribution aligned with the major axis of the host dark matter halo (e.g. Zentner et al. 2005; Zentner 2006; Knebe et al. 2004; Aubert et al. 2004; Faltenbacher et al. 2007, 2008). Thus by studying the orientation of satellite galaxies with respect to the major axis of the host *light profile* it may be possible to learn about the relative orientation between the stars in the host galaxy and the dark matter halo they reside in. In the Milky Way

and Andromeda, satellite galaxies appear in an anisotropic distribution, preferentially aligned with the minor axis of their hosts (e.g. Metz et al. 2009). When averaged, however, studies of satellites in SDSS find that satellites are randomly located around late-type galaxies (Bailin et al. 2008; Yegorova et al. 2011) and aligned with the major axis of the host light profile for early-type hosts (Brainerd 2005; Agustsson & Brainerd 2010). The studies of the satellite angular distribution outside of the Local Group have all been limited to relatively bright satellites (within 1-2 magnitudes of the host magnitude), and they have not attempted to measure the relative ellipticity of the satellite angular distribution. As stated above, measuring the angular distribution of fainter satellites is of key interest because several simulations predict that satellites that have very elongated orbits will lose the most gas as they pass through the host galaxy.

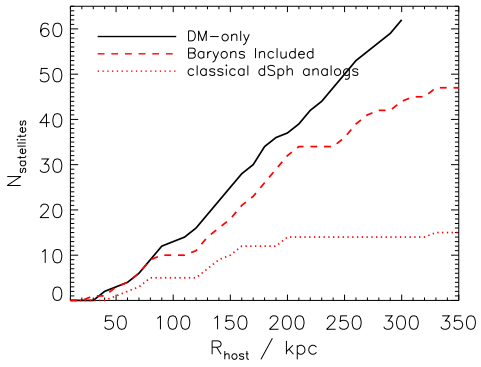

Figure 1.2: Theoretical model of the effects of baryons on the satellite radial distribution from a high resolution, hydrodynamical simulation by Brooks & Zolotov (2014). Both interactions between satellites and the host galaxy, as well as internal baryonic processes lead to a steeper prediction for the satellite radial number density relative to the dark matter only case.

In Chapters 2 and 3 we present new measurements of the satellite luminosity function using deep, high resolution HST imaging, which enable the detection of faint satellites which are up to a thousand times fainter than their hosts, at intermediate redshifts. In

Chapter 4 we demonstrate how these measurements provide important leverage in distinguishing between various models of star formation in satellite galaxies which otherwise predicted consistent results for low redshift satellites of Milky Way mass host galaxies.

1.2 Gravitational lensing

Strong gravitational lensing provides a powerful means of detecting dark subhalos at cosmological distances. Here we provide a brief overview of the method, assuming, without deriving key results from general relativity. For a derivation of these results we suggest the review by Carroll (2004). A more thorough discussion of the applications of strong lensing beyond measuring the subhalo mass function can be found in Treu (2010).

One of the major results of general relativity is that mass warps space-time, which in turn alters the geodesics that light follows. Figure 1.3 is a schematic representation of this effect. The viewer sees an image of the source which is offset from where the source would appear if there was no intervening lens. It is convenient to consider the 'lens equation':

From general relativity, the amount that the light is deflected in the direction perpendicular to the unperturbed path $\boldsymbol{\alpha}$ is given by an integral of the transverse gradient relative to the light travel path of the gravitational potential, Φ along the light travel path:

$$\boldsymbol{\alpha} = \frac{2}{c^2} \int \nabla_{\perp} \Phi(s) \; ds \tag{1.1}$$

The gravitational influence of the lens galaxy is often relevant only over a very short

distance scale relative to the total path of the light, in which case the 'thin lens' approximation can be used. In this approximation, the mass of the lens is projected onto a plane at the center of the three dimensional lens (i.e. the yellow ellipsoid in Figure 1.3).

Defining the two dimensional mass distribution projected along the z direction (from the observer to the lens) to be:

$$\Sigma(\boldsymbol{\xi}) = \int \Phi(\boldsymbol{\xi}, z')dz', \tag{1.2}$$

where $\boldsymbol{\xi}$ is the two dimensional vector in the plane of the lens. The gravitational potential in the thin lens approximation (Φ_t) becomes:

$$\Phi_t(\boldsymbol{\xi}) = 2G \int \Sigma(\boldsymbol{\xi'}) \frac{|\boldsymbol{\xi} - \boldsymbol{\xi'}|}{(\boldsymbol{\xi} - \boldsymbol{\xi'})^2} d\boldsymbol{\xi'} \tag{1.3}$$

The integral in equation 1.1 becomes a delta function in the z direction evaluated in the plane of the lens so that:

$$\boldsymbol{\alpha}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \Sigma(\boldsymbol{\xi'}) \frac{|\boldsymbol{\xi} - \boldsymbol{\xi'}|}{(\boldsymbol{\xi} - \boldsymbol{\xi'})^2} d\boldsymbol{\xi'} \tag{1.4}$$

It is useful to consider quantities relevant to the observer. Going back to Figure 1.3, we see that the 'true' transverse position of the source, $\boldsymbol{\beta}$ is related to the observed position of the source $\boldsymbol{\theta}$ simply by:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{LS} D_L}{D_S} \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1.5}$$

where the angular transverse position $\boldsymbol{\theta}$ is related to the physical transverse distance $\boldsymbol{\xi}$ by the angular diameter distance:

$$\boldsymbol{\theta} = \frac{D_L}{D_S} \boldsymbol{\xi} \tag{1.6}$$

The observed image position is thus determined by the gravitational potential of the deflector, and angular diameter distances which depend on cosmology. This result can be generalized to situations in which there are multiple significant lens galaxies at large z separations by studying multiple lens planes (see e.g. McCully et al. 2014a).

From Equation 1.4 it can be seen how gravitational lensing can be used to detect subhalos, via their alteration of the gravitational potential of a halo, which leads to a shifted observed image position relative to where the image would have appeared if the potential had been smooth. This method requires that the unperturbed gravitational potential of the lens is well determined at the position of the images, which is possible in the case of strong gravitational lensing.

In strong gravitational lensing, Equation 1.5 has multiple solutions, and thus multiple images of the background source appear. Each lensed image yields a constraint on the source position as well as the mass distribution of the deflector, known as the macromodel. If the lensed images are extended (as in the case of a lensed galaxy source), then there is significant angular constraint on the macro model, which makes it possible to robustly detect substructure via astrometric perturbations to the image positions. This method, known as gravitational imaging, proposed by Koopmans (2005), has been successfully applied to HST and Adaptive Optics (AO) imaging Vegetti et al. (2012, 2010a), and simulations indicate it will also work for ALMA data (Hezaveh et al. 2013). Recently Vegetti et al. (2014) analyzed a sample of 11 lenses, finding a fraction of substructure consistent with numerical simulations and results from analyses of four image quasar lenses.

The mass sensitivity of gravitational imaging is determined by the signal to noise

ratio and resolution of the data, as well as the intrinsic morphology of the lensed galaxy. Current measurements with Keck AO are sensitive to substructure with masses of $\sim 10^{7.5}\mathrm{M}_{\odot}$ and larger, while measurements with the Next Generation Adaptive Optics on Keck, the Thirty Meter Telescope, Very Long Baseline Interferomety may allow the sensitivity of this method to reach masses of $\sim 10^{6}\mathrm{M}_{\odot}$.

Gravitational imaging is a relatively new method for detecting perturbing substructure as it requires very high resolution imaging and intensive computation in order to rigorously model both the mass distribution of the deflector and the light profile of the lens source.

Traditionally, substructure has been detected via perturbations to the *magnification* of compact sources. The magnification is sensitive to lower mass perturbations as it depends on the second derivative of the potential, rather than the first derivative as in the case of the image position.

This can be seen by considering that the magnification is the ratio between the final and initial size of an area element after gravitational lensing occurs. To calculate the lensed size, we apply the differential lensing equation to see that the i^{th} component of the differential source element will be mapped onto the j^{th} component of the image according to the derivative of the lensing equation:

$$\frac{\partial \beta_i}{\partial \theta_j} = \frac{\partial}{\partial \theta_j}\left(\theta_i - \frac{D_{LS}D_L}{D_S}\alpha_i\right) = \delta_{ij} - \frac{D_{LS}D_L}{D_S}\frac{\partial \alpha_i}{\partial \theta_j} \tag{1.7}$$

The determinant of the above matrix gives the ratio of the initial size to the final size, and the magnification is the inverse of this.

As in the case of gravitational imaging, substructure is detected by deviations relative

to the smooth model prediction. When magnifications are used to detect substructure, the method is traditionally referred to as studying 'flux ratio anomalies', although this term is somewhat misleading given that the image magnifications provide a constraint on the local presence of substructure whether or not there is an anomaly.

In order to obtain robust constraints on the presence of substructure there are a few observational requirements. First, given that the image magnifications depend on the second derivative of the potential, the long range effect of a perturbing subhalo is smaller than in the case of gravitational imaging, thus deviations to the image magnifications are most apparent in lensed compact sources. Secondly, stars in the plane of the lens galaxy can induce high magnifications over small areas, thus the background source must be significantly larger than the lensing scale of these stars which is typically of order microarcseconds. Finally, as in the case of gravitational imaging, a tight constraint on the macro model is necessary in order to determine what the image magnifications should have been in the absence of substructure.

Traditionally, these requirements have been met by studying the image magnifications of the seven known quadruply imaged, radio loud, quasar sources. The radio jets are compact enough to be significantly magnified by substructure which has a characteristic lensing scale of order milliarcseconds, but they are thought to be extended enough to be unaffected by microlensing. The four lensed images as well as a constraint on the position of the host galaxy typically from optical imaging, have been shown to yield a sufficient constraint on the macro model to detect significant flux ratio anomalies induced by substructure. Dalal & Kochanek (2002) used a lensing simulation with smooth halos and perturbing subhalos to model the image magnification in these seven systems and inferred

that the lenses contained a fraction of mass in substructure which was broadly consistent with predictions from CDM simulations, albeit with large uncertainties. Progress with this technique has been limited by the small number of known suitable systems.

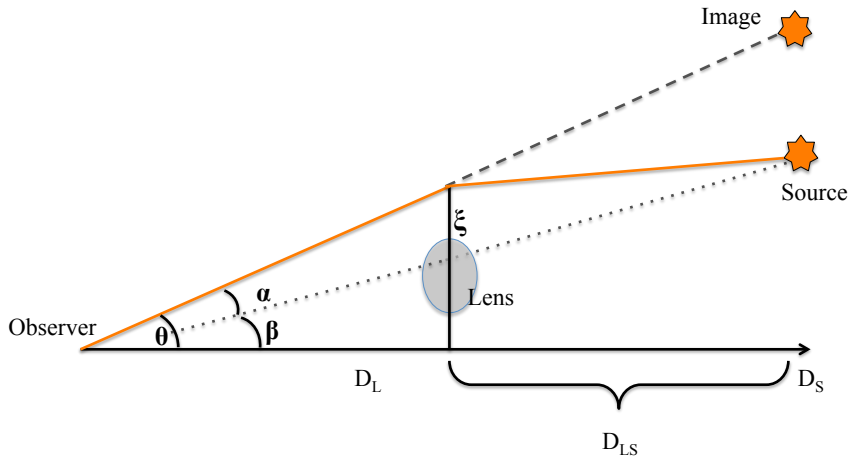

Figure 1.3: A schematic of the gravitational lensing effect. The path of light is curved due to the presence of the intervening lens galaxy in the image plane. The observer sees the light at a position θ which is a transverse distance α from β, the position the image would have been at if there had been no intervening galaxy. D_L, D_S and $D_{L,S}$ are the angular diameter distances from the observer to the lens, from the observer to the source and from the lens to the source respectively, and are relevant to gravitational lensing calculations.

In Chapter 6 we present a new method of using strong gravitational lensing to detect substructure, which will enable substructure to be detected in a much larger sample of systems than was previously possible.

Ultimately, measurements of the properties of luminous satellites and the subhalo mass function can be combined to provide complementary constraints on star formation physics in satellite galaxies. At present, the combination of gravitational lensing analyses with luminosity function measurements is the only method to directly measure stellar to virial mass ratios for satellites outside of the Local Group. In order to make such a comparison, it is necessary to compare satellite populations of lens and non lens galaxies

in an appropriate way. In Chapter 5 we compare the luminous satellite populations of lens and non lens galaxies, to test for systematic differences in the two.

Chapter 2

Luminous satellites I: Determining the number and spatial distribution in a Bayesian framework

This chapter was published as Nierenberg, A. M., Auger, M. W., Treu, T., Marshall, P. J., Fassnacht, C. D., "Luminous Satellites of Early-Type Galaxies I: Spatial Distribution", ApJ, 731, 44, (2011) and is included here with minor formatting adjustments

A powerful approach to understanding star formation efficiency in low mass satellite galaxies is the combination of lensing studies to constrain the mass function of satellites, and imaging studies to constrain the luminosity function (Treu 2010; Kravtsov 2010, and references therein). With this goal in mind, we have started a new program to characterize the visible properties of faint satellites of massive galaxies. In this first chapter, we focus on the spatial distribution of faint satellites of early-type galaxies at intermediate redshifts, $0.1 < z < 0.8$, selected from the GOODS fields (Giavalisco et al. 2004). We concentrate on this population of hosts because early-type galaxies dominate the sample of strong lensing galaxies (Auger et al. 2009) and are therefore the proper host sample for comparison to the satellite mass function results from lensing studies. An additional benefit to studying early-type galaxies is that they have relatively smooth surface brightness profiles which are ideal when searching for nearby compact and faint

companions.

For this initial analysis we only require one photometric band. We use z_{850} because it is the reddest of the GOODS bands and therefore it is the most faithful tracer of stellar mass.

This chapter is organized as follows: In §2.1 we discuss the properties of our host galaxy sample. In §2.2 we summarize our image analysis method, including our elliptical B-spline host galaxy subtraction and faint object detection and photometry methodologies. In §2.3 we take a first look at the satellite distribution by means of a binned analysis, which is useful for visualizing the main trends and identifying the strength of the signal. In §2.4 we describe our model for the combined satellite plus background object spatial distribution and the parameters that we aim to constrain (average number of satellites, power law slope, etc.) along with a number of nuisance parameters (density of the background population, slope of the background number counts, etc.). In § 2.5 we present the results obtained by comparing our model to the data. In § 2.6 we discuss our results and compare them to previous satellite studies. In § 2.7 we provide a concise summary. The Appendix contains more detailed explanations of many of the methods we used in this chapter.

2.1 Host Galaxy Sample

We select a population of early-type (E and S0) host galaxies from the catalog of Bundy et al. (2005), which contains spectroscopic redshifts,stellar mass estimates, and morphological classifications for 47% of its objects (see Treu et al. 2005a). The objects in

this catalog were originally selected from Hubble Space Telescope photometric catalogs [1] made by the GOODS team using the SExtractor software (Bertin & Arnouts 1996). In the GOODS-South field, we use COMBO-17 photometric redshifts (Wolf et al. 2004) for our hosts where spectroscopic redshifts are not available. In the GOODS-North field only a handful of objects have morphological classifications and stellar mass estimates but no spectroscopic redshifts. We excluded these from our analysis.

We limit the host redshifts to $z < 0.8$ to guarantee that we can detect satellites with host-satellite luminosity contrasts fainter than the equivalent contrast between the Small Magellanic Cloud and the Milky Way at all redshifts, thus increasing the likelihood that we observe approximately one satellite per host. We also exclude $z < 0.1$ galaxies, which are few (owing to the small volume) and too extended in angular size to analyze in the same way as the more distant sample. Finally we exclude two hosts which appear to be undergoing major mergers as these have physical environments which are significantly distinct from the majority of our sample. To ensure that we study the same population of satellites for all hosts, despite their broad distribution in redshift, we only consider satellites brighter than a fixed fraction (Δm) of the host galaxy luminosity (i.e. a fixed difference in magnitude). We also require that all objects be brighter than the detection threshold ($z_{850} = m_{\rm max} = 26.5$, as described in Section 2.2.2), and therefore we cut the parent sample to a maximum value of z_{850}, depending on the choice of Δm, such that $m_{\rm host} + \Delta m < m_{\rm max}$.

As we increase the size of the magnitude range we study, the number of hosts that are complete within that magnitude range drops; there are 202, 127 and 71 hosts complete

[1] The GOODS catalogs are available at `http://archive.stsci.edu/prepds/goods`

to Δm = 6.0, 5.5, and 5.0 respectively in the final host sample. At the same time, the number of satellites per host is expected to increase with Δm. The optimal choice of Δm needs to strike a balance between the two effects. As we will show, for the present GOODS dataset we find that Δm=5.5 maximizes the signal to noise ratio of the detection and therefore we will adopt this choice as our default. To illustrate the robustness of our results to small changes in Δm, we will also describe our findings for Δm = 5 and 6. The distributions of host redshifts, absolute magnitudes, and stellar masses for the three choices of Δm are shown in Figure 2.1. As Δm becomes larger, the hosts that satisfy the completeness requirements tend to become brighter and shift to lower redshift. As with any flux limited sample, the lower redshift objects of the sample will be dominated by the more abundant, intrinsically fainter galaxies, while the higher redshift objects will be dominated by intrinsically brighter objects. We defer an analysis of evolutionary trends to Chapter 3 Results in this chapter are an average of the properties of the satellite population in the $0.1 < z < 0.8$ redshift range.

We use the second order moment of the host intensity distribution, measured along the major axis by SExtractor [2] as a scale factor which we represent by $R_{\rm h}$. The use of $R_{\rm h}$ compensates not only for varying angular size with redshift, but is also intrinsically related to host masses via the size-mass relation (e.g. Trujillo et al. 2006; Williams et al. 2010) and thus adjusts for variations in host masses at a given redshift. $R_{\rm h}$ varies across the host sample in physical size (0.8-5.5 kpc) with host mass variations and in angular size (between 0.″2-1.″1). The median (and modal) values of $R_{\rm h}$ are 0.″5 and 3 kpc. We use these as fiducial values when converting our results to angular and physical scales.

[2] SExtractor 's AWIN_IMAGE

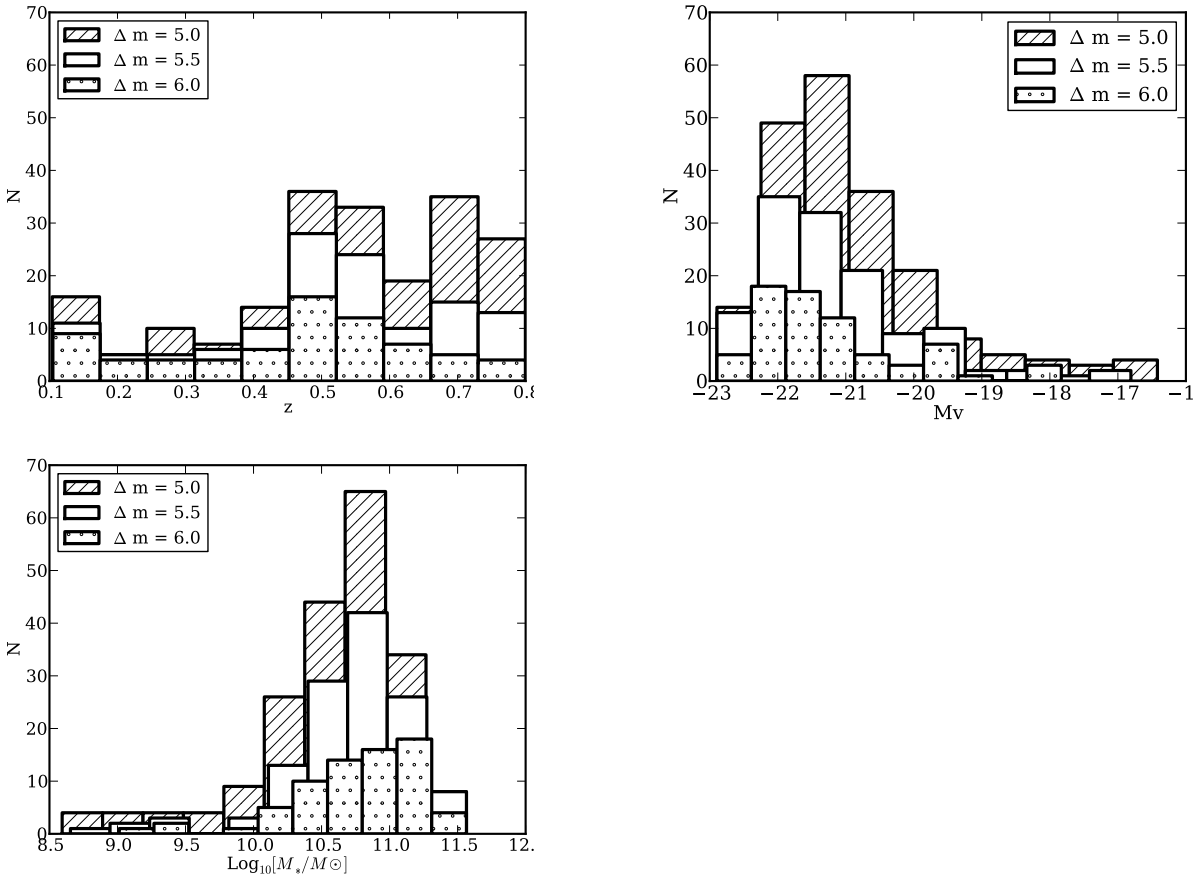

Figure 2.1: Clockwise from upper left: The distribution of host redshifts, absolute V band magnitudes, and stellar masses in three host completeness ranges ($\Delta m < m_{\rm max} - m_{\rm host}$) from Bundy et al. (2005).

2.2 Detection and photometry of close neighbors

Companions of high redshift galaxies are difficult to study because they are intrinsically faint and often obscured by the host galaxy light. In this section we describe our method of modeling and subtracting the host galaxy light profile to allow us to identify and perform accurate photometry on nearby objects. Note that while we model host light in small regions around each host, we do not limit our analysis to objects within this modeled region. We use the GOODS catalogs for photometry and astrometry for all objects outside of the modeled cutout region.

2.2.1 Host Galaxy B-Spline Models

Producing a good model for the light profile of the host galaxies is critical for studying the population of faint objects near the host. We use a multi-step process to create an accurate model of the host light profile which does not include the light from nearby objects.

We model the host light profile in a small cutout centered on the host galaxy. We choose this cutout to be 20×20 $R_{\rm h}$ in order to ensure that all significant levels of host light are removed. In the first step of the modeling process, we use SExtractor to identify all objects in the cutout region. The SExtractor parameters at this stage are chosen to err on the side of identifying noise peaks as objects in order to ensure that all real objects being obscured by the host light are recognized; see Appendix 2.8.1 for a detailed description of the SExtractor parameters and how they were chosen. SExtractor outputs a segmentation map which we use to mask out all identified objects other than the host galaxy. SExtractor also returns measurements of the axis ratio and position angle of the host galaxy which we use in addition to $R_{\rm h}$ to define an elliptical coordinate system for the cutout.

We then model the host light in the masked image using empirical polar B-spline models. We choose B-spline models because they quickly fit the observed light distribution with a smooth model that is independent of PSF effects and is more flexible than Sersic models, for example. Our method is similar to that described by Bolton et al. (2005, 2006). Each iteration of the B-spline code fits a model to the masked data, subtracts the model, and then identifies and masks new residual structures. This process

is repeated three times. The final B-spline model is then subtracted from the data to produce a residual image, which is used to perform object detection and photometry for field galaxies near the host. Further details of the modeling and masking procedure are provided in Appendix 2.8.1. Examples of our host subtraction procedure are shown in Figure 2.2.

2.2.2 Object Detection and Photometry

We use SExtractor to detect and measure the properties of objects within the host-subtracted cutouts. To ensure completeness, we limit the sample to objects with MAG_AUTO $z_{850} < 26.5$ magnitudes, where the GOODS images are virtually 100% complete, even close to the host galaxy as we will show below. Furthermore, we only study objects fainter than $z_{850} > 21.0$ magnitudes. This is because our analysis relies on an accurate characterization of the background (see Section 2.4.2) and very bright objects appear in low numbers with large fluctuations which can bias the number counts slope near a particular host.

Because we use the GOODS catalog data outside of our cutouts, it is imperative that our detections and photometry after host light subtraction are as close as possible to GOODS detections and photometry. We confirmed this by running SExtractor with our parameters on a large un-modeled section of the GOODS field. We recovered virtually the same number of objects and our photometry was consistent with GOODS photometry. This comparison is described in further detail in Appendix 2.8.2.

Due to the relative brightness of hosts compared to satellites, there is a minimum

radius at which we will be able to accurately identify faint objects, regardless of how careful we are during the host modeling process.

By simulating point sources at and fainter than our chosen limiting magnitude (26.5) at varying distances from a representative subset of hosts, we find the minimum radius for completeness to be 1.5 $R_{\rm h}$ ($\sim 0\farcs9$) for the vast majority of cases. The few cases for which highly flattened hosts have significant residuals extending to a larger distance are identified manually; for those systems, appropriately larger inner regions are excluded from our analysis (see Appendix 2.8.1).

Note that in the redshift range we study, even the most intrinsically faint sources will have effective radii typically less than $0\farcs1$ (see, e.g., de Rijcke et al. 2009) which is smaller than the FWHM of the GOODS PSF. Thus these sources are effectively point sources. We test our sensitivity to intrinsically faint sources by simulating faint exponential disks near our hosts at varying redshifts, with effective radii estimated from the relations given by de Rijcke et al. (2009). In the innermost region, between 1.5 and 3.5 $R_{\rm h}$, we failed to detect approximately 20% [3] of 26.5 magnitude objects. Outside of this region we recovered approximately 100% of the simulated objects. For brighter objects with apparent magnitudes of 24.5 our recovery was approximately 100% complete even in this inner region. In our final analysis we study the satellite population between 1.5 and 45 $R_{\rm h}$. Taking this into account, we can examine the worst case scenario in which all satellites are exponential disks with apparent magnitudes of 26.5. Assuming that satellites are distributed radially in projection as $P_{sat}(r) \propto r^{-1}$, we would only underestimate the

[3]The true incompleteness estimate requires knowledge of the object luminosity function. We make a rough estimate by averaging the results for objects at redshifts of 0.1, 0.4 and 0.8, where the redshift 0.1 objects are the most intrinsically faint and most difficult to detect.

final satellite number by 1-2 % because even in this extreme case, the incompleteness is confined to a small region. Thus we expect that the loss of completeness in the innermost region for the faintest sources will have a negligible effect on the analysis of the total sample.

To compile a single catalog and avoid duplications, we compare the position of each object identified in the residual images to objects already in the GOODS catalogs. If the object is already in the catalogs (position within $0\farcs3$), we replace the GOODS photometry and astrometry with our own measurements, which do not suffer from being contaminated by host-galaxy light (see Figure 2.3). The mean distance between 'matched' objects was 4 ± 3 pixels, this corresponds to $0\farcs12$ which is the FWHM of the GOODS PSF. For objects that are not matched to objects already in the GOODS catalog, we add a new entry. Finally, for all objects outside of the host-subtracted cutouts, we use the measurements from the GOODS catalog and we do not attempt to detect new objects. In Figure 2.2 we show examples of the modeling process for a variety of host ellipticities and physical sizes, along with newly detected objects in the host-subtracted images.

2.2.3 Objects Detected in Cutout Regions

The host light subtraction has two important effects on our measurement of objects near the host galaxy. The first is that it removes host light contamination and allows for accurate photometry of nearby objects, as shown in Figure 2.3. This figure compares GOODS photometry to host-subtracted photometry for objects that had already been identified in the GOODS catalogs. As expected, we measure object magnitudes to be

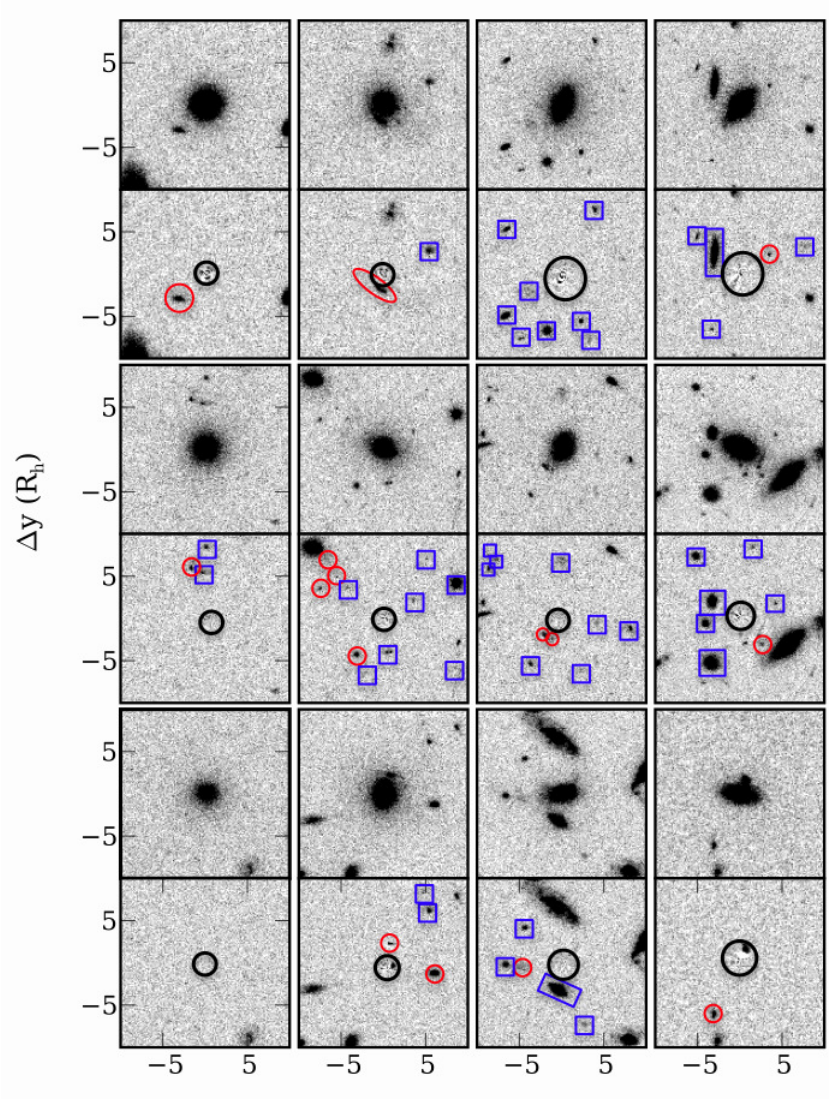

Figure 2.2: Demonstration of galaxy modeling for a range of host ellipticities and sizes. Original images are shown with residuals after host model subtraction immediately below. Host galaxy physical size and ellipticity decrease downwards and to the left. All images are 20 $R_{\rm h}$ on a side. New object detections, made possible by the host light subtraction, are circled in red. Objects that were detected in the GOODS catalogs and that have complete photometry within the cutout (i.e. did not raise a SExtractor flag greater than 2) are identified by blue squares. Some objects visible to the eye were omitted from our final catalog because they were fainter than our detection limit (for instance, the object just outside of the excluded region in the top right galaxy). The central region of the host, excluded in our analysis, is identified by a black circle.

slightly fainter than the GOODS photometry after host subtraction, with a mean difference of 0.12 ± 0.02 magnitudes for objects within 8 $R_{\rm h}$ ($\sim 4\farcs$) of the host galaxy. Our photometry is identical to GOODS photometry without host subtraction (see Appendix 2.8.2), and we confirmed that the host light removal was accurate by simulating faint point sources near the hosts and ensuring that they were recovered with accurate photometry. Thus the difference in magnitude after host subtraction shown in Figure 2.3 is entirely due to the removal of the host light contamination which has a significant impact on photometry performed near bright objects.

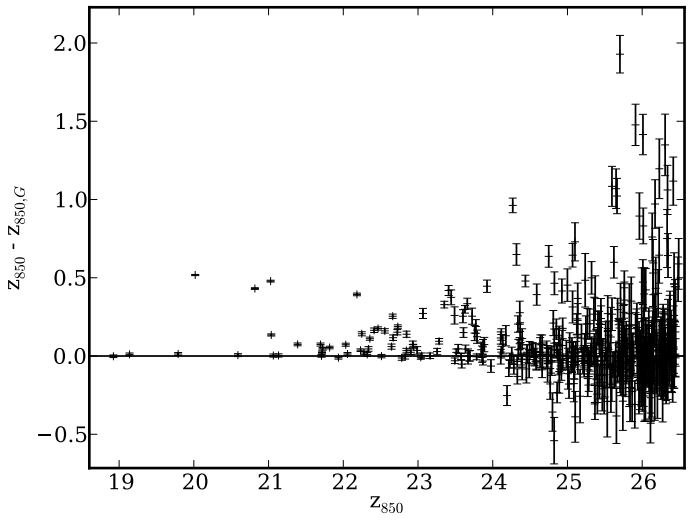

Figure 2.3: Comparison of photometry before ($z_{850,G}$) and after (z_{850}) host subtraction for objects in cutout regions which had already been identified in the GOODS catalogs prior to host subtraction.

The second important effect of host light subtraction is that it allows us to detect new objects. If the newly detected objects are real rather than artifacts of the host light subtraction, then we expect that object properties such as brightness and elongation will be similar to the properties of objects that had already been detected in the GOODS fields. The effect of weak lensing on magnification and shear is negligible for the small

number of background sources close to the host Einstein radii in projection (the affected region is typically an annulus of $\sim 1\farcs0 \pm 0\farcs2$ for massive ellipticals). One way to see this is using the fact that the effect of weak lensing on the background number counts goes as $N_{\rm obs}/N_{\rm true} = 1/\mu^{(\beta-1)}$ where μ is the magnification due to the host galaxy and β is the power-law slope of the faint end of the background galaxy flux distribution (Schneider et al. 2006, see Equation 111 in Part 1) In the case of z_{850}, β is fairly close to one (about 0.7) and thus we do not expect weak lensing to have a significant impact on our object detection.

In the top panel of Figure 2.4, we compare the distributions of the axis ratios of the GOODS catalog sources and the newly detected objects near the host galaxies. The two distributions are indistinguishable, with a Komogorov-Smirnoff (KS) probability of being drawn from the same distribution of 0.96. The second panel of Figure 2.4 shows the distribution of contrast in MAG_AUTO (δm$= m - m_{\rm h}$) between host and detected objects. Note the use of lower-case δ, which denotes a specific contrast from the host and is different from Δm which describes the allowed maximum contrast between host and neighboring objects for a particular data set. The KS value for the two distributions being the same is 0.14. This value is low, but not low enough to rule out the possibility that the distributions are the same. Note that our improved host galaxy light subtraction procedure is important for detecting companion objects fainter than about δm$=$ 2.5 magnitudes (which is about 0.5 magnitudes fainter than the magnitude contrast between the Milky Way and the Large Magellanic Cloud).

As expected, the number density radial profiles are very different for newly detected objects and objects that had already been detected in GOODS (bottom panel of Fig-

ure 2.4), with a KS probability of $6.0 \cdot 10^{-25}$ for new and GOODS objects being drawn from the same spatial distribution. The number density of new detections increases steadily towards the center of the host, while the number density of objects already in the GOODS catalogs decreases. Host light subtraction triples the number density of detected objects in the inner 3″.

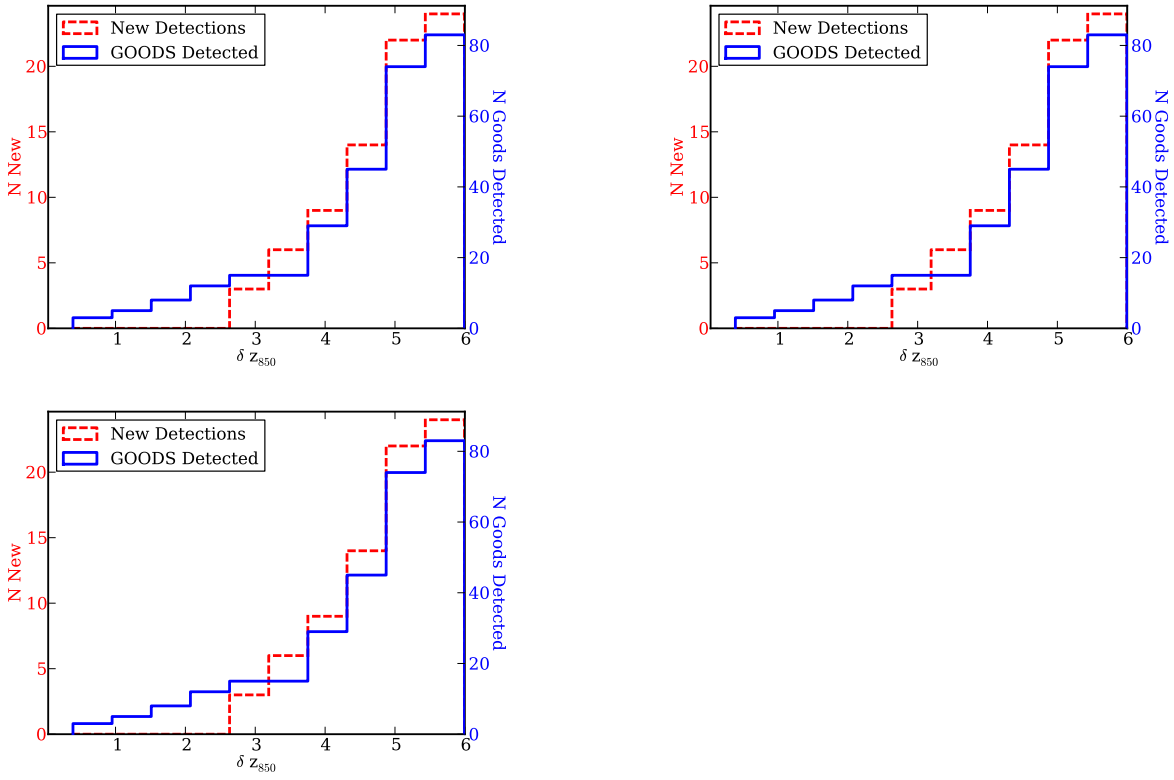

Figure 2.4: Comparison of the properties of objects detected in the GOODS catalogs to those of newly detected objects. *Top Left*: The distribution of object ellipticities ($1 - q = 1 - b/a$). *Top Right*: The distribution of magnitude differences from hosts (δm= $m - m_h$). *Bottom*: Number density of objects as a function of distance from the host. Newly detected objects are closer to the host than those in the GOODS catalogs.

2.3 First Look

The host light subtraction discussed in Section 2.2 allows us to create an enhanced catalog of objects near the host galaxies. In this section, we show the radial and angular profiles of objects in spatial bins in order to provide a qualitative sense of the properties of objects near the hosts. Binning is useful because it provides a visual representation of data. However, it is inherently limited because it requires data points to be averaged, thereby losing valuable information. Thus we do not perform our analysis on the binned data but instead use this section to justify our model choices in Section 2.4.

2.3.1 Radial Distribution

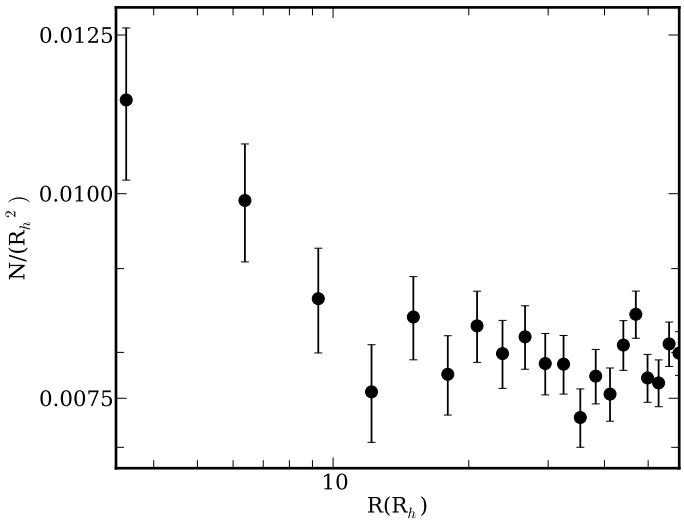

Figure 2.5: Average number density of objects as a function of distance from the host in units of the second order moment of the host intensity profile along the major axis ($R_{\rm h}$).

In Figure 2.5 we show the average number density of objects as a function of distance from the hosts. The number density of sources increases dramatically near the hosts.

At large radii the number density becomes dominated by the isotropic and homogeneous distribution of objects not associated with the hosts. In Section 2.4 we will describe how we analyze the number density signal by inferring the combined properties of the satellite and background/foreground populations.

2.3.2 Angular Distribution

In Figure 2.6 we plot the angular distribution of objects within 20 $R_{\rm h}$ of the host galaxies, where $\phi = 0$ is aligned with the host major axes. We show the distribution of $|\phi|$ only for hosts with q$<$ 0.6 in order to ensure all object angles are well measured (for round hosts it becomes more difficult to measure the host position angle). The figure shows that objects appear with more frequency towards $\phi = 0$ than would be expected for a uniform distribution (shown by a dashed line). We also compare the observed angular distribution of objects to a uniform distribution by applying a KS test which rules out the angular distribution being uniform with 95 % confidence. Recall that this is *without* any kind of attempt to separate the background signal from the satellite signal. We investigate this asymmetry more rigorously in Section 2.4.

2.4 Joint Modeling of Satellite and Background Galaxy Populations

We have shown that the neighboring galaxies to the GOODS host galaxies exhibit non-trivial radial and angular distributions, and we now seek to model these distributions.

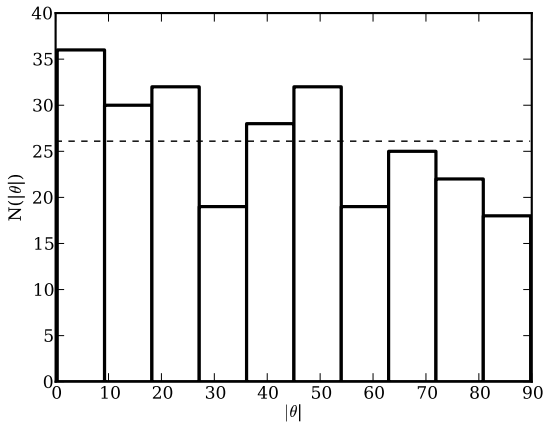

Figure 2.6: Number of objects at angle θ where $\theta = 0$ is aligned with the host major axis within radius of 20 $R_{\rm h}$ for hosts with axis ratio less than 0.6. The dashed line shows the average bin heights for a uniform distribution.

We start in § 2.4.1 by defining the satellite model and the parameters we aim to constrain. In § 2.4.2 we discuss the background model and the priors we will be using. In § 2.4.3 we summarize our inference methodology, which is described in more detail in Appendix 2.9. A summary of all model parameters and their priors is given in Table 2.1. Our implementation of the model and inference algorithm has been extensively tested by means of simulated data. To guide the reader, a schematic of a possible realization of our satellite plus background model is shown in Figure 2.7.

2.4.1 Satellites

The scale-free nature of ΛCDM results motivates us to construct our model for the spatial distribution of the satellite number density as a function of corresponding host parameters. For simplicity, and owing to the relatively low signal-to-noise ratio of our measurement, we do not allow for intrinsic scatter in these scaling functions. Thus our inferences connecting host and satellite properties are averages for the entire population;

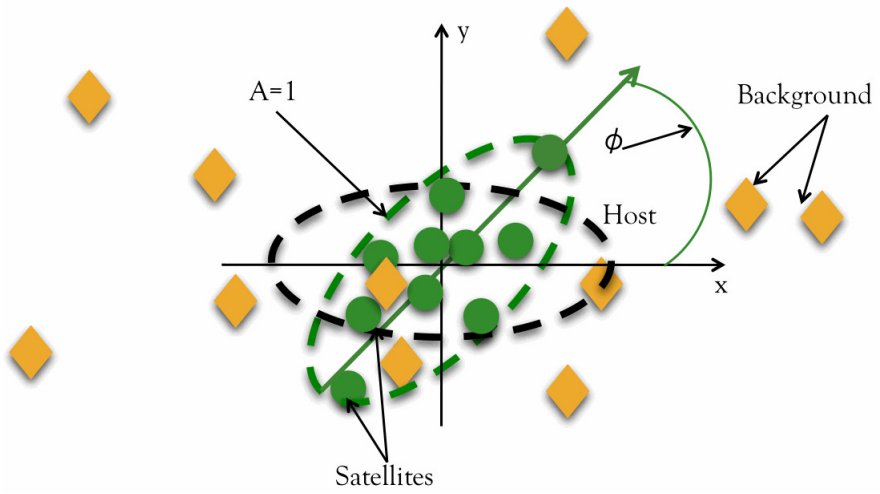

Figure 2.7: Schematic illustrating a possible realization of our model.

individual systems will differ from the average.

We divide our discussion of our model for the satellite population into three subsections. We describe the probability of finding a satellite at a certain position in Subsection 2.4.1. We discuss choosing a region in which to search for satellites in order to normalize the spatial distribution in Subsection 2.4.1. Finally the probability of finding a certain number of objects around each host is discussed in Subsection 2.4.1.

Satellite Spatial Distribution

ΛCDM simulations indicate that the number density of satellites follows the mass density profile of their host galaxy (Kravtsov 2010). In turn, observations indicate that the three dimensional total mass density profile of elliptical galaxies can be approximated by a power law $\rho^{-\gamma'}$ with $\gamma' \approx 2$, (with <10% scatter) (e.g. Koopmans et al. 2009a; Auger et al. 2010). The power law profile seems to extend as far as 100 $R_{\rm e}$ (*e.g.* Gavazzi et al. 2007; Lagattuta et al. 2010) which corresponds to approximately 70 $R_{\rm h}$. We will thus

model the radial distribution of satellites as a power law, with projected logarithmic slope $\gamma_{\rm p} = 1 - \gamma'$.

We relate the angular light profile of the hosts to the mass profile following the results from the Sloan Lens ACS (SLACS) by Bolton et al. (2008, hereafter B08). B08 used gravitational lensing to study the relationship between the mass and light distributions of massive elliptical galaxies. B08 found that the major axes of the light profiles of the galaxies they studied were well aligned with the total central mass profiles and that the axis ratios of the light and mass profiles were the same within measurement errors; note that this agreement was established within the Einstein radii of the lensed galaxies, which corresponds to roughly $\sim 1''$ (see also Kochanek 2002). The majority of our hosts are well represented by the properties of the SLACS lenses which are massive elliptical galaxies with stellar masses in the range $\sim 10^{10.5}$ to $10^{12} M\odot$ (Auger et al. 2009)[4]. This implies that we can expect the angular distribution of satellites to roughly follow that of host light and therefore that it is reasonable to construct a model of the satellite angular profile which is related to parameters which describe the host light angular profile.

We model the satellite distribution in an elliptical coordinate system described by an angle θ which is the same as the normal polar angle, and a radial coordinate R' related to Cartesian coordinates x, y by:

$$R' = \sqrt{x^2 + y^2/q^2} \tag{2.1}$$

Where q is the ratio between minor and major axes of the elliptical coordinate system. The probability of finding a satellite on an elliptical contour within the area element

[4]We assume that the small number of hosts (~ 10) in our sample which were less massive do not deviate significantly from the relationship between mass and light orientation observed in their more massive counterparts.

$R'dR'd\theta'$, goes as a power law with power $\gamma_{\rm p}$

$$P_{\rm s}(R')R'dR'd\theta' \propto (x^2 + y^2/q^2)^{\gamma_{\rm p}/2}dxdy. \tag{2.2}$$

Changing to polar coordinates gives

$$P_{\rm s}(r, \theta|\phi, q, \gamma_{\rm p}) \propto$$

$$r^{\gamma_{\rm p}}[\cos^2(\theta - \phi) + 1/q^2 \sin^2(\theta - \phi)]^{\gamma_{\rm p}/2} r dr d\theta \tag{2.3}$$

where we introduce ϕ to allow for some offset between the major axis of the satellite elliptical coordinate system and the major axis of the host galaxy; $\phi = 0^o$ corresponds to parallel alignment, and $\phi = 90^o$ corresponds to perpendicular alignment (see Figure 2.7). This is different from the coordinate θ, which describes the offset between a particular object and the host major axis. We are only interested in the magnitude of the offset between the host light profile and the satellite population so we infer $|\phi|$, the absolute value of the offset of the distribution from the host major axis.

In general the offset ϕ in 3 dimensions is related to the 2D projected offset in a non-trivial way because of the random orientation of the host-satellite system in the plane perpendicular to the projection plane. However, for many interesting and plausible scenarios, the alignment in 3 dimensions relates in an obvious way to the observed, projected offset ϕ. Namely, if the satellites are aligned with the host light distribution in 3D then the projection of the two systems will also appear aligned. Similarly, if the systems are anti-aligned, the satellites will appear perpendicular to the host in projection. Finally, if the satellites are oriented randomly with respect to the host in 3 dimensions, their projection will appear isotropic. Thus the projected relationship between the satellite

spatial distribution and the host light profile contains relevant information about the 3D distribution.

We also aim to determine the connection between the ellipticity of the host light and that of the satellite distribution. For this purpose, we define the ellipticity to be:

$$\epsilon = \frac{1 - q^2}{1 + q^2}, \tag{2.4}$$

We introduce a parameter A to relate the ellipticity of the host profile, $\epsilon_{\rm h}$, to that of the satellite distribution, $\epsilon_{\rm s}$:

$$\epsilon_{\rm s} = \frac{\epsilon_{\rm h} A}{1 + \epsilon_{\rm h}(A - 1)} \tag{2.5}$$

We choose this parametrization because it returns valid ellipticities ($0 < \epsilon < 1$) for all values of $A > 0$. This is convenient when exploring the A parameter space with our MCMC code. The parameter A can be understood as follows: when $A = 0$, the satellite distribution is always round regardless of how flat the host distribution is; when $A = 1$ the satellite distribution has the same flattening as the host distribution; and for values of $A > 1$, the satellite distribution is more flat than the host light distribution. Note that formally A goes from zero to infinity. As A approaches infinity, the satellite distribution approaches $q = 0$. We are interested in a qualitative characterization of the satellite flattening so we simplify our inference on A by restricting our analysis to $0 < A < 2$, which will allow us to distinguish between isotropic and flattened satellite distributions without having to explore an unnecessarily large parameter space.

Spatial Normalization

We choose a maximum and minimum radius in which to search for satellites in order to normalize our distribution. These radii are determined by observational constraints. The inner radius is constrained to be $r_{\rm min} = 1.5R_{\rm h}$ by our ability to accurately measure faint object magnitudes near the host (see Section 2.2.2). For the outer radius we choose $r_{\rm max} = 45R_{\rm h}$ which corresponds to approximately 140 kpc. This choice is a compromise between studying an area large enough to find LMC/SMC equivalent objects (which are $\sim$ 60 kpc from the Milky Way), and keeping the area small enough to limit overlap between host galaxies and allowing us to reasonably apply a background prior determined by the entire field.

The normalized radial probability distribution is thus:

$$P_{\rm s}(r|\gamma_{\rm p}) = \left(\frac{\gamma_{\rm p}+2}{r_{\rm max}^{\gamma_{\rm p}+2} - r_{\rm min}^{\gamma_{\rm p}+2}}\right) r^{\gamma_{\rm p}+1}, \tag{2.6}$$

The normalization of the angular distribution is given by a generalized elliptical integral.

As discussed in Section 2.2.2, a few of our systems are not complete to the same inner radius due to issues with light modeling for highly flattened hosts. Furthermore, some of the regions far from the hosts are not complete to 45 $R_{\rm h}$ due to GOODS field edge effects. We discuss how we account for this incompleteness in Appendix 2.8.3.

Number of satellites per host

Multiple dark matter simulations have found that the number of satellites with a given mass relative to the host mass is constant (Moore et al. 1999; Kravtsov et al. 2004b; Gao et al. 2004). With this in mind, we model the number of satellites within

a fixed *magnitude* range from the host as being drawn from a Poisson distribution with some constant mean $N_{\rm s}$.

It is important to keep in mind the following caveats. Due to baryonic physics, the observable properties of galaxies are not exactly scale invariant. In fact, the virial mass to light ratio is not a universal function and we expect it to vary for the host galaxies and satellites. This means that in a fixed magnitude range we are not actually probing a fixed virial mass range, but only a fixed range in luminosity ratio. However, it should be noted that our host galaxies are typically luminous enough (several $10^{10} L_{\odot}$) that their satellites are also significantly brighter than the typical Local Group dwarf galaxies where the M/L is believed to be much higher than in massive ellipticals. Thus our satellite mass range is better characterized by the relative luminosity between host and satellites than it would be for the Local Group.

Naturally, these assumptions will have no effect on our inference of the parameters of the satellite spatial distribution, provided that it is also independent of host galaxy luminosity within the spatial and mass ranges considered here.

2.4.2 Background/Foreground Objects

In addition to satellites, each of the host galaxies is surrounded in projection by background/foreground objects. We isolate the satellite signal by using the properties of the entire GOODS fields to provide a strong constraint on the background number density.

Galaxies tend to cluster on scales of order a typical galaxy virial radius, or ~ 400 kpc

(e.g. Totsuji & Kihara 1969; Peebles 1974; Brainerd et al. 1995; Villumsen et al. 1997; Zehavi et al. 2002; Morganson & Blandford 2009). This clustering is believed to be due to the accretion of matter along dark matter filaments (e.g. Benson 2010, and references therein). Because of clustering, the density of objects tends to be higher in regions near bright galaxies and to have fluctuation amplitudes larger than Poissonian. Chen et al. (2006) tested a variety of methods for removing 'interloper' (background/foreground) contamination from their satellite signal using a set of ΛCDM simulations. Chen et al. (2006) found that simply estimating the number density of background objects by studying randomly selected regions in their field significantly underestimated the contamination from foreground/background objects that appeared near their hosts in projection. They found the most reliable estimate of the background was obtained by measuring the background in annuli centered on their hosts and just outside of the area in which they studied the satellite population. We adopt this method to build a prior on the background near the hosts. We study the background in annuli between 45 and 60 $R_{\rm h}$ (typically 140-180 kpc). We choose this distance range to ensure that we are not including satellites in our estimate, while still accurately characterizing the clustering of galaxies near the hosts. We find that the density of objects near the hosts with magnitudes within our detection range ($21 < z_{850} < 26.5$) is $\Sigma_{b,o} = 125 \pm 2$ arcmin^{-2}. As expected, this is higher than the average density of objects in the GOODS fields (117 ± 0.6) and has larger fluctuations than one would predict from Poisson noise.

Recall that we are studying objects brighter than a fixed magnitude contrast from the host magnitudes. This means that the number of objects we study around a given host is a fraction of the density of objects brighter than $z_{850} = 26.5$. We correct for this

around each of the hosts by representing the cumulative distribution function (CDF) of the background number counts by a power-law (e.g. Benítez et al. 2004)

$$N_b(< m_{\rm max}) \propto 10^{\alpha_{\rm b} m_{\rm max}} \tag{2.7}$$

where $\alpha_{\rm b}$ is the slope of the background number counts. The maximum magnitude $m_{\rm max}$ for a particular host system is

$$m_{\rm max} = m_{\rm h} + \Delta m. \tag{2.8}$$

Thus for a given choice of Δm, the expected number density of background/foreground objects around the $j^{\rm th}$ host is

$$\Sigma_{\rm b,j} = \Sigma_{\rm b,o} 10^{\alpha_{\rm b}(m_{\rm h,j} + \Delta m - m_{\rm lim})}. \tag{2.9}$$

We measure the background slope to be $\alpha_{\rm b} = 0.28 \pm 0.01$ in the same annuli near the hosts in which we estimate the number density of the background.

2.4.3 Analysis

In the previous two subsections we constructed a model which describes the probability of a satellite or foreground/background object appearing a given distance from the host dependent on a choice of parameters. The parameters and their priors are listed in Table 2.1. Of key importance in our work is that our results are unbinned and analyzed in a fully Bayesian fashion. This means that for each parameter our result is a posterior probability distribution function (PDF) which describes the probability of a value of a parameter being true given our data (the likelihood) and our prior knowledge of the parameter (the prior).

Parameter	Description	Prior
Satellite Model		
Ns	Number of satellites per host	U(0,20)
$\gamma_{\rm p}$	Logarithmic slope of the satellite radial distribution	U(-5,0)
A	Flattening of the satellite number density distribution relative to host light flattening	U(0,2)
$\|\phi\|$	Offset between the major axis of satellite spatial distribution and that of the host.	U(0,π/2)
Background Model		
$\Sigma_{\rm b,o}$	Number density of background objects with magnitudes between $21 < z_{850} < 26.5$	N(125,2)
α_b	Logarithmic slope of the z_{850} background number counts	N(0.28,0.01)

Table 2.1: Summary of model parameters. U(a,b) denotes a uniform distribution between a and b, and N(μ,σ) denotes a normal distribution with mean μ and standard deviation σ.

In Appendix 2.9 we discuss the details of constructing the posterior probability function (Equation 2.17) which we use to study our parameter posterior PDFs. We compute the posterior PDFs using a Markov Chain Monte Carlo (MCMC) method. At least 10^6 iterations per chain are performed in order to ensure convergence.

2.5 Results

We first discuss the results for the $\Delta m = 5.5$ magnitude bin, which has the strongest signal. Table 2.2 contains a summary of the results of the inference. The posterior PDF for each variable is shown in Figure 2.8.

The first main result is the clear detection of a peak in the posterior PDF for $N_{\rm s}$ which is well removed from zero, indicating the detection of a population of satellites. Secondly, we find that the radial density profile of the satellite spatial distribution is consistent with isothermal ($\gamma_{\rm p} = -1$). Thirdly, the posterior PDF of ϕ is peaked at zero indicating that the satellite spatial distribution is preferentially aligned with that of the host. A KS test shows that a uniform distribution of angles is ruled out at more than a 99.99% CL.

Our results are inconclusive for the relationship between ellipticity of the satellite distribution and that of the host, described by the parameter A (see Figure 2.8). This is not surprising as it takes a stronger signal to measure the ellipticity of a distribution than to measure its alignment. Our inability to infer A is in part due to a degeneracy between ϕ and A. If A is zero, all values of ϕ are equally probable. This can be seen in the contour plot shown in Figure 2.8. We show the effects of removing this degeneracy by performing a separate analysis on the $\Delta m = 5.5$ data set, keeping ϕ fixed at zero. The results for this are shown in Figure 2.9. When ϕ is fixed at zero, the inference disfavors small values of A, which implies a disfavoring of satellite distributions that are rounder that that of the stars.

We also verified that our results are robust to small changes in limiting magnitude by repeating the analysis for $\Delta m = 6.0$ and 5.0 magnitude bins. The $\Delta m = 6.0$ bin clearly shows the presence of a satellite population, with the same isothermal slope observed for the $\Delta m = 5.5$ range and the same angular alignment with the host major axis. The inferred satellite number is also consistent with the $\Delta m = 5.5$ measurement, although there is a longer tail towards higher satellite numbers which might indicate a slightly higher number, as expected given the fainter limit. As discussed in Section 2.1, the number of hosts that are complete almost halves from $\Delta m = 5.5$ to 6.0 so the errors are larger for the $\Delta m = 6.0$ inference overall. Interestingly, the analysis for $\Delta m = 5.0$ did not conclusively detect a satellite population; the number of satellites is marginally more than 1-σ greater than zero, and the uncertainties on the parameters describing the spatial distribution are similarly larger. This shows the crucial need for deep data in performing this kind of measurement. In all cases we recover our priors on the background

parameters α_b and $\Sigma_{\rm b,o}$. We are not able to constrain these numbers further because in our model they are both degenerate with each other and with the number of satellites.

An important degeneracy is that between $N_{\rm s}$ and $\gamma_{\rm p}$. The inferred number of satellites is larger for a shallow radial profile and vice-versa. From theoretical and observational arguments (see the Introduction) we expect the radial profile to be close to isothermal and we expect alignment between host and satellites. It is thus instructive to repeat the analysis by fixing these parameters to their expected values $\gamma_{\rm p} = -1$ and $\phi = 0$, thereby eliminating many of the degeneracies. The resulting posterior PDFs for $\Delta m = 5.5$ are shown in Figure 2.10, while results for all magnitude ranges are summarized in Table 2.3.

As expected, fixing $\gamma_{\rm p}$ and ϕ lowers the uncertainty of our inference and the detection of satellites becomes more significant. It is also easier to compare the results across magnitude bins, and we see how the number of satellites indeed increases with Δm and is consistent with a cumulative luminosity function going as L^{-1} (i.e. similar to the cumulative mass function predicted from theory assuming $m \propto L^{-2}$), albeit with large errors. Furthermore, the posterior PDF of A begins to disfavor a satellite population that is more isotropic than the stars in the galaxy. This is also broadly in line with expectations, as we expect that the distribution of stars has been made rounder than the host halo by dissipational processes.

2.6 Discussion

This is the first measurement of the numbers and spatial distribution of faint satellites of early-type galaxies at intermediate redshift so it is useful to provide some context for

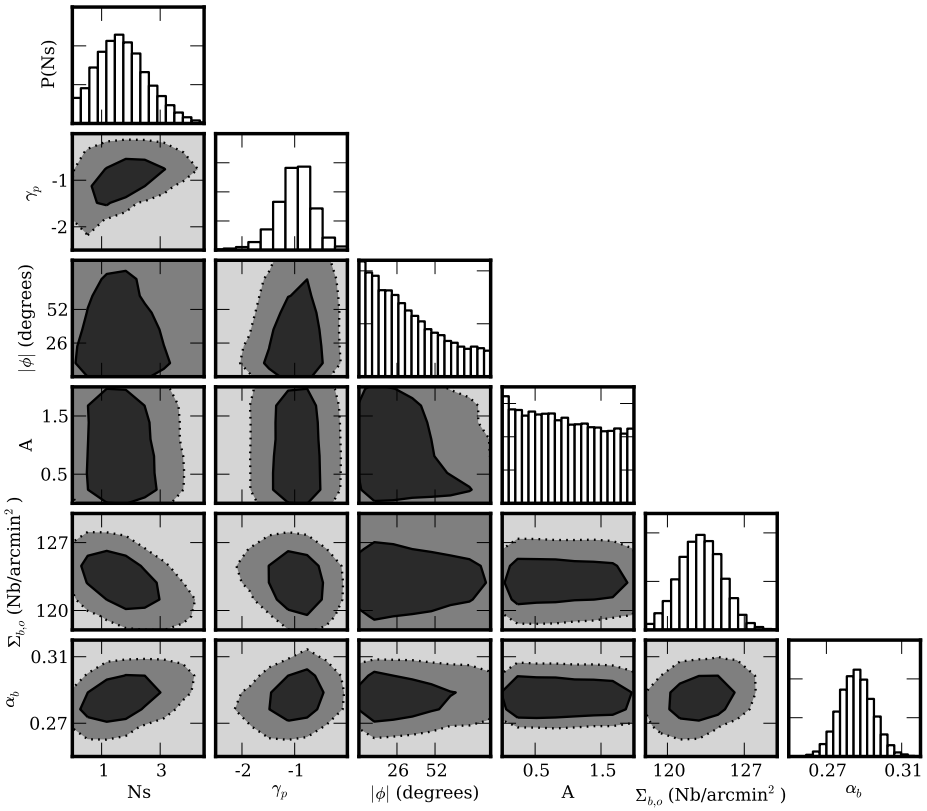

Figure 2.8: Bivariate posterior PDFs for all model parameters for the $\Delta m = 5.5$ data set. The dark and light contours contain regions of 68 and 95% of the probability respectively. The diagonal shows the marginalized PDF for each parameter.

our result. Our measurement of the radial slope of the satellite spatial distribution is consistent with an isothermal distribution, i.e. $\gamma_{\rm p} \approx -1$. This result is similar to that found in W10, and the radial distribution of satellites appears to be consistent with that of the total mass distribution measured in lensing studies (e.g., Koopmans et al. 2009a; Auger et al. 2010). However, given the uncertainty on the inferred slope, the spatial distribution of satellites is also consistent with the NFW profile inferred by C08. In the radial range covered by our work ($\sim$5–140 kpc), an NFW profile is characterized by

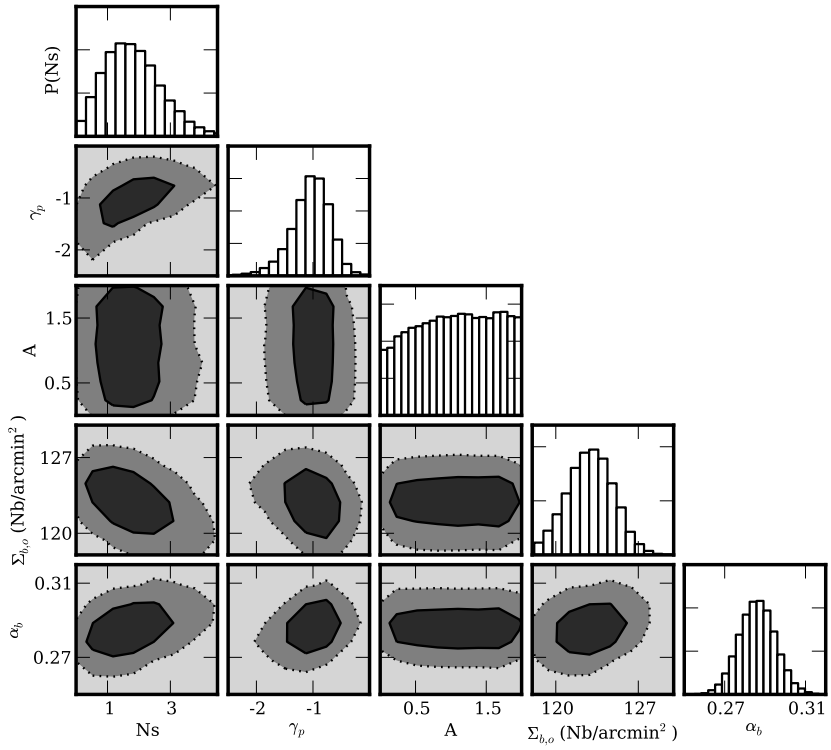

Figure 2.9: Same as Figure 2.8, but with ϕ fixed at 0.

a radially averaged projected logarithmic slope of approximately $\gamma_{\rm p} \sim -1$, with large variations depending on host scale radius. More accurate measurements of the radial density profile of satellites are needed, in addition to a comparison of the radial profile around different host masses, in order to determine whether the distribution of satellites follows that of the total mass or that of the dark matter component.

The distribution of satellites is anisotropic and is preferentially aligned along the major axis of the host light profile. This host-satellite alignment is consistent with observations of host-satellite systems in SDSS (e.g. Brainerd 2005). Furthermore, alignment is

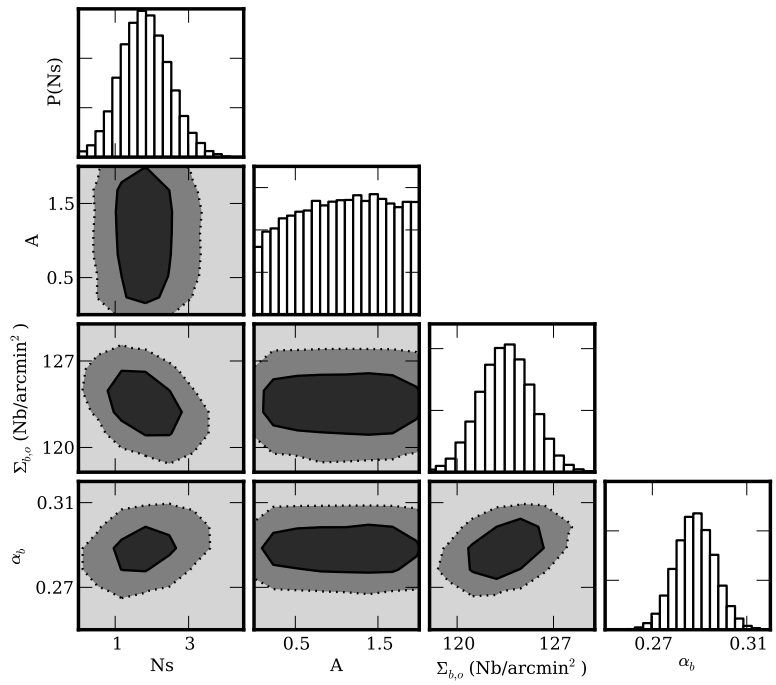

Figure 2.10: Same as Figure 2.8, but with fixed $\phi = 0$ and $\gamma_{\rm p} = -1$.

predicted by ΛCDM simulations which show satellites accreting anisotropically along filaments and appearing aligned with the major axis of the host mass profile (e.g., Aubert et al. 2004). Previous observations have established that the host light profile aligns with the host mass profile (B08), and our observation of the satellite-host light alignment therefore implies alignment between the satellite distribution and the host mass, in agreement with simulations. This alignment between host mass and satellite spatial distribution has important implications for the frequency of flux ratio anomalies, as we discuss below.

Δm	$\gamma_{\rm p}$	Ns	$\lvert\phi\rvert$ (68% confidence)	A	$\Sigma_{\rm b,o}(\rm N_b/arcmin^2)$	α_b
6.0	$-1.1^{+0.5}_{-0.6}$	3^{+1}_{-1}	< 44	- (a)	123^{+2}_{-2}	$0.283^{+0.009}_{-0.009}$
5.5	$-1.0^{+0.3}_{-0.4}$	$1.7^{+0.9}_{-0.8}$	< 42	-	123^{+2}_{-2}	$0.286^{+0.009}_{-0.009}$
5.0	$-0.5^{+0.4}_{-0.8}$	$0.5^{+0.8}_{-0.4}$	< 56	-	124^{+2}_{-2}	$0.282^{+0.009}_{-0.008}$

Table 2.2: Posterior Medians/Confidence Intervals, (a) No inference on the parameter because the posterior distribution is approximately uniform

Fixed parameter	Δm	$\gamma_{\rm p}$	Ns	A (68% confidence)	$\Sigma_{\rm b,o}(\rm N_b/arcmin^2)$	α_b
$\phi = 0$	5.5	$-1.0^{+0.3}_{-0.4}$	$1.7^{+1.0}_{-0.8}$	> 0.73	123^{+2}_{-2}	$0.286^{+.009}_{-0.009}$
$\phi = 0$, $\gamma_{\rm p} = -1$	6.0	-	2^{+1}_{-1}	> 0.83	124^{+2}_{-2}	$0.285^{+0.009}_{-0.009}$
$\phi = 0$, $\gamma_{\rm p} = -1$	5.5	-	$1.8^{+0.7}_{-0.6}$	> 0.72	124^{+2}_{-2}	$0.288^{+0.008}_{-0.008}$
$\phi = 0$, $\gamma_{\rm p} = -1$	5.0	-	$0.4^{+0.4}_{-0.3}$	> 0.73	124^{+2}_{-2}	$0.281^{+0.008}_{-0.008}$

Table 2.3: Posterior Medians/Confidence Intervals with fixed parameters

Although we do not expect the number of satellites of intermediate redshift elliptical galaxies to be exactly equal to that of galaxies at lower redshift due to baryonic processes and evolution, it is instructive to compare the GOODS satellites to other satellite populations. Figure 2.11 compares the cumulative luminosity function (CLF) of our satellites to the CLF of the Milky Way and Andromeda satellites (adopted from Tollerud et al. 2008), with one sigma uncertainties calculated from Gehrels (1986), and to the CLF of SDSS satellites of hosts with varying morphologies adopted from C08. Ideally, we would have also liked to compare our study to J10 which is one of the few studies of satellites of high redshift objects. However, the maximum value of Δm in that survey was about 2.5 magnitudes which is too bright to be compared to our measurement in a meaningful way.

C08 measured the number of satellites and their radial profile between approximately

20 and 350 kpc. In order to compare our results appropriately, we use the C08 'interloper subtracted' fit to a radial power law given in their Table 2 in order to extrapolate their measurement inward toward the host. We assume that the power law is constant in the inner regions and use a fiducial value of $R_{\rm h} \sim 3$ kpc to estimate the number of satellites C08 would have seen in the region that we studied (i.e., as close as 1.5 $R_{\rm h}$). The Milky Way CLF is complete in the same equivalent region as ours so we did not have to make any adjustments in order to compare our numbers. All three studies measured magnitudes in different filters; C08 used r band magnitudes at an effective redshift of $z = 0.1$, we use observed-frame z_{850} magnitudes, and Milky Way measurements are in rest-frame V. However, the corrections required to convert from r at redshift ~ 0.1 and observed z_{850} at redshifts of ~ 0.5 into rest-frame V are negligible compared to the sizes of the bins of Δm that we study, and we therefore do not include explicit k-corrections in our analysis. The three satellite CLFs are roughly consistent where the measurements overlap, given the relatively large error bars (Figure 2.11). It is worth noting that our GOODS satellite measurements have significantly smaller error bars than those for the Local Group by virtue of the large sample of hosts we study which allows us to reduce the sampling error. Furthermore, we are able to observe much fainter satellites than the SDSS study.

It is also instructive to compare our inferred satellite number to the number of minor mergers our hosts are expected to undergo in the time span we study. As many of our satellites are very near their host galaxies, we expect that some of them will be close to merging and that our estimate of the satellite number should account for the predicted minor merger rate in the mass range we study. The merger rate depends on

the ratio between the host and satellite virial masses (e.g., Fakhouri et al. 2010). Our typical host has a stellar mass of M* $\sim 10^{10.5}$ M$\odot$ (see Figure 2.1). From abundance matching techniques (e.g., Behroozi et al. 2010), this corresponds to a halo mass of about 10^{12} M$\odot$. In the luminosity range of our satellites the stellar mass to light ratio of the satellites should be approximately in the range 30-100% of that of the host galaxy (e.g., Kauffmann et al. 2003). Using this, the stellar mass of the satellites we study is in the range of $0.6-2\cdot10^{8}$ M$_{\odot}$. This corresponds to virial masses of $\sim 2-8\cdot10^{10}$ M$_{\odot}$, according to the best fit to Equation 21 in Behroozi et al. (2010). Thus the we can detect satellites with virial masses of order a percent of that of their hosts.

This estimate is subject to large uncertainties, including those arising from tidal forces that tend to remove dark matter from satellites as they move closer to their host galaxy, thus reducing their virial mass. Using the best fit to Equation 1 in Fakhouri et al. (2010), we estimate that the majority of our hosts (60%) will undergo about 0.4-0.5 mergers (depending on the satellite stellar mass to luminosity fraction) between redshifts 0.1 and 0.8. It is reassuring that this number can be easily accounted for by our inferred satellite number. Of course, a quantitative comparison requires a detailed evaluation of the merging and visibility timescales, which is beyond the scope of this Chapter. However, we will re-visit this point in a future work, as minor mergers have been suggested to play an important role in the evolution of early-type galaxies (e.g. Bundy et al. 2007; Boylan-Kolchin et al. 2008; Naab et al. 2009; Nipoti et al. 2009; Kaviraj et al. 2009; Hopkins et al. 2010; Kaviraj et al. 2011).

Finally we discuss the effects our detected satellite population would have on flux ratio anomalies in quadruply-imaged strong gravitational lens systems. We limit our

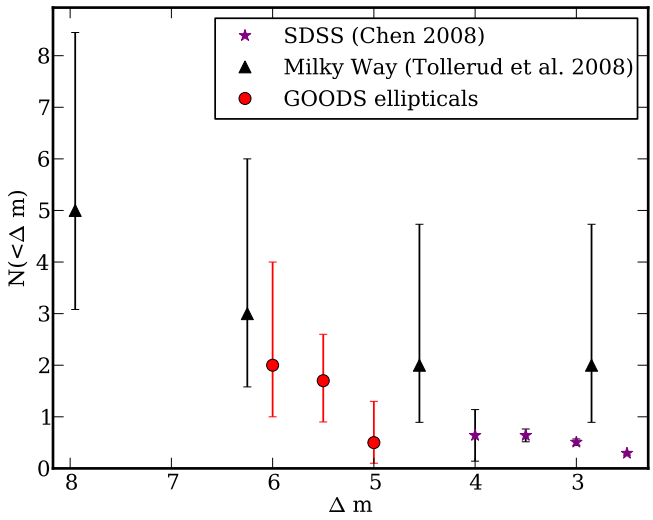

Figure 2.11: Comparison of the measured number of GOODS satellites to those found for the Milky Way and SDSS galaxies.

discussion here to order of magnitude estimates and general trends, leaving a rigorous quantitative analysis to a future paper.

Dalal & Kochanek (2002, hereafter D02) used the magnifications of images in five quad lenses to infer a mass profile for the lens galaxies. In order to reproduce the observed magnifications, D02 inferred the presence of a non-smooth component to the mass distribution with a mass fraction between 0.6 and 7 % near the lensed images. Mao et al. (2004) and more recently Xu et al. (2009, hereafter X09) used ΛCDM simulations to test whether dark substructure embedded within smooth halos could produce the flux ratio anomalies observed by D02. These studies estimated that the mass fraction of CDM substructure from their simulations would only be about 0.1 % near image positions, which is insufficient to produce the flux ratios studied by D02. The inclusion of globular cluster populations and baryonic processes does not change this result (Xu et al. 2010) and underscores the importance of constraining the satellite population through direct

observations.

From our inference, a few percent of our hosts have a satellite within the annulus that would include the typical lens image positions studied by X09 (the Einstein radius is approximately $1''$). We therefore find that the average mass fraction in this annulus is of order a few percent. Using the mass function ($dn/dm \propto m^{-1.9}$) described by X09, we estimate that the average satellite mass fraction they would have observed in our mass range is of order a percent, which is consistent with our result. It is worth considering the possibility that the presence of a few relatively massive satellites may increase the lensing cross-section significantly so that any attempt to measure the satellite mass function using gravitational lenses would be automatically biased towards very high mass fractions near the lens Einstein radius; we will explore this possibility in more detail in Chapter 5.

It is interesting to note that, although our typical hosts have similar halo masses to those studied in X09, the most massive subhalo X09 observed in any of their simulations has a virial mass of $\sim 10^{10}$ M $\odot$, which is on the very low mass end of the subhalos studied in this work. Our simple extension of the virial to stellar mass relationship led us to estimate that our hosts will have on average more than one halo *at least* this massive even assuming that the mass to light ratio of our satellites is only a few percent of that of the host halo. This discrepancy may be due to the fact that our mass estimate is very approximate and, as already discussed, may not be appropriate for systems undergoing strong interactions such as tidal stripping. On the other hand, the difference in estimated satellite masses may indicate the importance of baryonic condensation in preserving the mass in subhalos during accretion (see e.g., Weinberg et al. 2008; Romano-Díaz et al. 2009).

Our observation that satellites appear preferentially along the host major axis may offer an additional route to reduce or eliminate the apparent discrepancy in the satellite mass fraction as inferred by flux ratio anomalies. As pointed out by Zentner (2006), an anisotropic satellite distribution can increase the effective projected mass of the host 'felt' by a lensed image by as much as a factor of six. Thus, the mass associated with our observed luminous satellites could in principle be sufficient to explain the anomalies detected by D02.

Our satellites are much fainter than a typical lensed quasar image and thus would be easily obscured in any of these lensing studies. Our result highlights the importance of studying the satellite population in non-lens systems in conjunction with the analysis of perturbations in lens systems in order to attain an understanding of the mass and luminosity functions of satellites.

To conclude, it is important to point out that the largest discrepancies between the luminosity and mass function are seen for even fainter satellites than the ones probed by this study (Kravtsov 2010). It is thus desirable to push the analysis of both the luminosity and the mass function of satellites to even smaller fractions of the host galaxy properties.

2.7 Summary

We use an advanced host light subtraction method to study the spatial distribution of faint galaxies around 127 early-type galaxies in the GOODS fields between redshifts 0.1 and 0.8. We employ a self-consistent model in the framework of Bayesian inference

to disentangle the satellite population from background/foreground galaxies. Exploiting the depth and resolution of the HST images, we detect satellites up to 5.5 magnitudes fainter than the host galaxy and as close as $0\farcs5/2.5$ kpc to the host. Our main results can be summarized as follows:

1. Intermediate redshift, massive, early-type galaxies have on average $1.7^{+0.9}_{-0.8}$ satellites within our luminosity ratio limit. This is consistent with the number of satellites observed in the Milky Way.

2. The number density of satellites follows an approximately isothermal radial power law profile $P(r) \propto r^{\gamma_{\rm p}}$ with $\gamma_{\rm p} = -1.0^{+0.3}_{-0.4}$.

3. The satellites are preferentially aligned along the major axis of the host light profile with $|\phi| < 42^o$ at the 68% confidence level.

4. When the offset ϕ between the satellite population and host light profile is fixed at 0 degrees, the satellite distribution is inferred to be preferentially more elongated than the distribution of host galaxy light.

5. When the satellite number density profile is assumed to be isothermal, the average number of satellites is inferred to be $1.8^{+0.7}_{-0.6}$.

Acknowledgements

We thank N. Jackson, L.V.E. Koopmans, R. Wechsler, M. Busha, P. Schneider, D.D. Xu, for many insightful comments and stimulating conversations. We thank M. Giavalisco and the rest of the GOODS team for their work on the GOODS ACS images

and catalogs. We thank J. Chen for providing data for Figure 2.11. AMN and TT acknowledge support by the NSF through CAREER award NSF-0642621, and by the Packard Foundation through a Packard Fellowship. PJM was given support by the Kavli Foundations and the Royal Society in the form of research fellowships.

2.8 Appendix

2.8.1 Host Galaxy Light Subtraction

As we showed in Section 2.2, the detectability of satellite galaxies is quite sensitive to the presence of light from the much brighter host galaxy. In this appendix we give more details on how the light from the host galaxy was subtracted in order to allow the faint satellites to be detected. Host galaxy surface brightness modeling is a two step process. In the first step, we identify and mask all objects other than the host in the region in which we want to model host galaxy light. This ensures that our host model does not attempt to fit the light of nearby objects. Once the objects are identified and the mask is created, in the second step we model the host galaxy surface brightness in 2D using the unmasked parts of the image, interpolate over the masked parts, and subtract this model from our original image.

Masking

The masking is done automatically using the segmentation map produced by an initial SExtractor run, with parameters tuned to optimize the detection of small faint objects in the presence of the host light. This optimization was performed by simulating

Parameter	Large Object Mask	Point Object Mask	Final Photometry
DETECT_MINAREA	10	5	–
DEBLEND_NTHRESH	64	–	–
DEBLEND_MINCONT	0.001	–	–
DETECT_THRESH	–	2.5	–
ANALYSIS_THRESH	–	2.5	–
FILTER_NAME	–	gauss_2.0_3x3.conv	–
BACK_TYPE	MANUAL	MANUAL	MANUAL
BACK_VALUE	0.0	0.0	0.0

Table 2.4: SExtractor parameters. Parameters not listed, or marked as "–", are those used by the GOODS team and can be found at http://archive.stsci.edu/pub/hlsp/goods/catalog_r2.

faint companion objects at varying positions around our hosts. We found that using the GOODS SExtractor parameters at this stage gave unsatisfactory results, in that objects we simulated near the hosts were not being masked. We found that, in order to properly account for both faint point sources near our hosts and diffuse sources further from our hosts, we needed to use a superposition of masks created with two different sets of SExtractor parameters. These parameters are listed in Table 2.4. With this masking routine, we were able to accurately recover magnitude 27 point sources as close as 1.5 $R_{\rm h}$ (typically $\sim 0\farcs9$) from our host galaxy centroids.

B-spline model subtraction

In Figure 2.2 we show some representative examples of our host galaxy light subtraction process. Some of our hosts are very elongated or show disk-like features. The surface brightness distribution of these objects is more difficult to model and tends to leave residuals at larger radii than the rounder hosts. Residuals are easily identifiable because of the symmetric pattern and distinct elongated shape that they appear in (see,

e.g., the residuals of the upper right host in Figure 2.2). We deal with this small number ($\sim$ 20) of more disky/extended sources in two ways. The first way is to increase the amount of flexibility in the B-spline fit, increasing the number of multipoles fitted in each of our spline rings around the ellipse. The second thing we do is to exclude a larger inner region from our analysis to ensure we are not identifying any residuals as objects.

2.8.2 Source Extractor Parameters and Photometry Comparison

After subtracting the main galaxy light out of the image we run SExtractor on the residual image to identify the remaining objects. In this step we modify our SExtractor parameters by increasing the deblending threshold in order to match the parameters used when making the GOODS catalogs. A full list of our SExtractor parameters is listed in Table 2.4.

Even after matching the deblending parameters, there are two minor differences between our final SExtractor parameters and GOODS SExtractor parameters. The first is that GOODS estimates the background by measuring the noise in an annulus with 100 pixel ($3''$) width around each object. This method of background subtraction is not optimal for our measurement because we are focusing on a relatively small area (typical image size was $\sim$ 200x200 pixels), within which we expect to have a relatively high object density. Instead, we estimate the background by calculating the 3 σ clipped mean within our image. Furthermore, the GOODS team made modifications to SExtractor that changed the deblending process slightly.

We compare the MAG_AUTO output from SExtractor for the two methods in a 4 arcmin^2 cutout from the GOODS field in order to study the effects of the different background subtraction and deblending (see Figure 2.12). Using our parameters, we identify 484 objects with MAG_AUTO<26.5. Of these, 22 objects do not have a center within 0.″3 of an object in the GOODS catalog. The major outliers in the MAG_AUTO comparison are all in areas of high object density and thus most likely due to differences in deblending. The mean difference (not including major outliers) in the MAG_AUTO estimate is $(-1.9 \pm 6.0) \cdot 10^{-3}$.

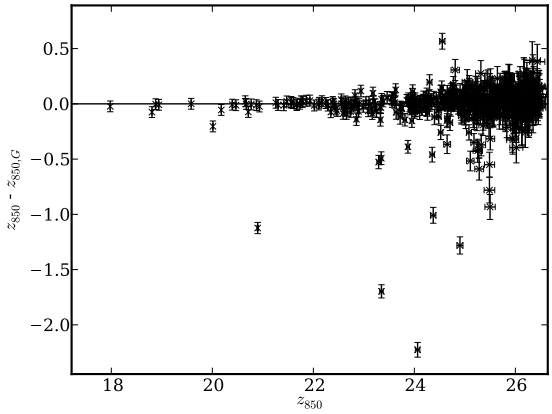

Figure 2.12: Comparison of our photometry with GOODS in a 4 arcmin^2 field cutout with no host galaxy subtraction.

2.8.3 Completeness

In Section 2.4 we discuss our model for the spatial distribution of satellites which we can use to calculate the probability that there are a certain number of objects $\hat{N}$ within $1.5 < r/R_h < 45$, with positions $\boldsymbol{x}$. Before comparing a particular model prediction to the data it is necessary to account for observational limitations such as field edge

effects. In this section we explain how we account for these observational limitations when inferring our model parameters.

Following Kelly (2007), we define a completeness function, $P(I|\boldsymbol{x})$, which describes the probability of observing an object at position $\boldsymbol{x}$, where $I = 1$ indicates an object is detected and $I = 0$ indicates a non-detection. Note that because we have chosen to study only magnitude ranges in which we are complete, our completeness function will only depend on position in a given field. As discussed in Appendix 2.8.1, we measure our completeness near the host galaxy by placing simulated objects around the host and calculating our recovery rate and the accuracy of our photometry as a function of position. Far from the host, we estimate our completeness in annuli to account for field edge effects. To simplify this analysis, we ignore regions of partial completeness so that a region in the field is either 100% complete or 0% complete.

Taking completeness into account, the likelihood of detecting $N^{\rm obs}$ total objects (satellites and background),with positions $\{\boldsymbol{x}\}$, given model parameters $\boldsymbol{\theta}$, becomes

$$\Pr(N^{\rm obs}, \{\boldsymbol{x}\}|\boldsymbol{\theta}) = \Pr(I|\{\boldsymbol{x}\})\Pr(\hat{N}, \{\boldsymbol{x}\}|\boldsymbol{\theta}) \tag{2.10}$$

Our fields have varying sizes but the model parameter $\hat{N}$ gives the number of objects we expect to find in an ideal field where we have 100% completeness between 1.5 and 45 $R_{\rm h}$. As discussed above, some of our host systems have restricted areas, between r_{mn} and r_{mx}, where the total area is smaller than in the ideal case. In these cases we must scale the model prediction by the probability of finding the object in the reduced area compared to the probability of finding the object in the ideal area, and we therefore

define an updated prediction for the number of objects $\hat{N}'$ given by

$$\hat{N}' = \hat{N} \frac{\int_{r_{mn}}^{r_{mx}} \Pr(\boldsymbol{x}'|\boldsymbol{\theta}) d\boldsymbol{x}'}{\int_{1.5}^{45} \Pr(\boldsymbol{x}'|\boldsymbol{\theta}) d\boldsymbol{x}'} \tag{2.11}$$

Note that $\hat{N}$ is a smooth model prediction of the number of true objects in a field. In order to properly account for the fact that the true number of objects in a field is discrete, we must introduce a Poisson probability, which relates the "true" discrete number of objects to the model number of objects. The likelihood that we observe $N^{\rm obs}$ objects at positions $\boldsymbol{x}$ given our model parameters, is a product over the likelihood of each of those positions being true given our model parameters

$$\Pr(\{\boldsymbol{x}\}|\boldsymbol{\theta}) = \prod_{i}^{N^{\rm obs}} \Pr(\boldsymbol{x}_i|\boldsymbol{\theta}) \tag{2.12}$$

Thus the probability that we observe $N^{\rm obs}$ given a set of model parameters, taking into account varying total areas of completeness, is given by:

$$\Pr(N^{\rm obs}, \{\boldsymbol{x}\}|\boldsymbol{\theta}) = \Pr(N^{\rm obs}|\hat{N}') \prod_{i}^{N^{\rm obs}} \Pr(\boldsymbol{x}_i|\boldsymbol{\theta}). \tag{2.13}$$

Properly, we should also include a term which accounts for the probability of *not observing* $N - N^{\rm obs}$ objects. However, in our case, the region of parameter space we exclude due to observational limitations is always much smaller than the region we do study, and thus the number of objects our model predicts that we expect to observe is much larger than the number that we cannot observe. Therefore, for simplicity, we do not include this term in our analysis.

2.9 Inference methodology

In this section we discuss in detail how we characterize the posterior PDFs for our satellite model parameters $\boldsymbol{\theta}_{\rm s} = \{N_s, \gamma_{\rm p}, A, \phi\}$, and our background model parameters $\boldsymbol{\theta}_{\rm b} = \{\Sigma_{\rm b,o}, \alpha_b\}$. We present a top-down description of our model in which we start with the posterior PDF we would like to obtain and break it apart to show where we have inserted different pieces of information.

Our final data set, $\boldsymbol{D}$, is composed of a set of measurements for all of our host-field systems, $\boldsymbol{D} = \{\boldsymbol{D}_1, \boldsymbol{D}_2, ..., \boldsymbol{D}_{N_h}\}$ Where $\boldsymbol{D}_j = \{\boldsymbol{h}_j, \boldsymbol{d}_j\}$, and $\boldsymbol{h}_j$ is the set of measurements of the magnitude, axis ratio and RMS of the j^{th} host light profile, and $\boldsymbol{d}_j = \left[N_j^{\rm obs}, \{(\boldsymbol{x}_1), (\boldsymbol{x}_2), ...(\boldsymbol{x}_{N_j^{\rm obs}})\}\right]$ is the number of objects observed around the j^{th} host and their positions. Using Bayes' theorem we can express the probability of a set of model parameters being true given the data by:

$$\Pr(\boldsymbol{\theta}|\boldsymbol{D}) \propto \Pr(\boldsymbol{D}|\boldsymbol{\theta})\Pr(\boldsymbol{\theta}) \tag{2.14}$$

In the above equation the term $\Pr(\boldsymbol{D}|\boldsymbol{\theta})$ is the likelihood of the data being true given a set of model parameters and $\Pr(\boldsymbol{\theta})$ is the prior information we have about the model parameters. Our model parameters and their priors are listed in Table 2.1.

The first term on the right of Equation 2.14 is a product of likelihoods for individual host systems:

$$\Pr(\boldsymbol{D}|\boldsymbol{\theta}) = \prod_j^{N_{\rm h}} \Pr(\boldsymbol{d}_j|\boldsymbol{\theta}, \boldsymbol{h}_j) \tag{2.15}$$

Notice that in our model, the measurements of the host light profile are treated as model parameters which we assume are known exactly. This is because we have

constructed our model such that its parameters do not predict the properties of the host light profile.

Equation 2.15 is itself composed of a product of the likelihood of measuring $N_j^{\rm obs}$ objects around each host times the probability for measuring each object position,

$$\Pr(\boldsymbol{d}_j|\boldsymbol{\theta},\boldsymbol{h}_j) = \Pr(N_j^{\rm obs}|\boldsymbol{\theta}) \prod_i \Pr(\boldsymbol{x}_i|\boldsymbol{\theta},\boldsymbol{h}_j) \tag{2.16}$$

Recall that in Section 2.4 we built a model that was composed of separate contributions from a satellite and background population. This means that the probability of finding any object at a given location is the sum of the probability of finding a background object at that location with the probability of a finding a satellite at that location:

$$\Pr(\boldsymbol{d}_j|\boldsymbol{\theta},\boldsymbol{h}_j) = \Pr(N^{\rm obs}|\boldsymbol{\theta}) \prod_i \Pr(\boldsymbol{x}_i|\boldsymbol{\theta},\boldsymbol{h}_j,S)\Pr(S|\boldsymbol{\theta},\boldsymbol{h}_j) + \Pr(\boldsymbol{x}_i|\boldsymbol{\theta},\boldsymbol{h}_j,B)\Pr(B|\boldsymbol{\theta},\boldsymbol{h}_j) \tag{2.17}$$

The term $\Pr(\boldsymbol{x}_i|\boldsymbol{\theta},\boldsymbol{h}_j,S)$ is the probability of finding a satellite at a certain position (given by Equation 2.3) and $\Pr(\boldsymbol{x}_i|\boldsymbol{\theta},\boldsymbol{h}_j,B)$ is the probability of finding a background/foreground object at a certain position (which is uniform). The term $\Pr(S|\boldsymbol{\theta},\boldsymbol{h}_j)$ is the probability of an object being a satellite given only the model parameters,

$$\Pr(S|\boldsymbol{\theta},\boldsymbol{h}_j) = \frac{N_s}{N_s+N_b} \tag{2.18}$$

and $\Pr(B|\boldsymbol{\theta},\boldsymbol{h}_j)$ is defined in an analogous way for the line-of-sight interlopers. The functions that determine N_s and N_b in terms of model parameters are discussed in Section 2.4.

Chapter 3

Luminous satellites II: Cosmic evolution and dependence on host properties of the satellite luminosity function and spatial distribution

This chapter was published as Nierenberg, A. M., Auger, M. W., Treu, T., Marshall, P. J., Fassnacht, C. D., Busha, M. T., "Luminous Satellites II Spatial Distribution, Luminosity Function and Cosmic Evolution", ApJ, 752, 99, (2012) and is included here with minor formatting adjustments

In this chapter we address three questions: 1) What is the spatial distribution and cumulative luminosity function of the satellites of massive galaxies? 2) How does the spatial distribution vary with satellite luminosity? 3) How do these properties vary with host morphology, stellar mass and redshift? We conclude by discussing how our results can be interpreted in the context of ΛCDM.

In the previous chapter, we developed a method of modeling and subtracting the host light profile to allow us to detect faint ($m_{sat} - m_{host} > 5.5$) satellites at intermediate redshifts near early-type host galaxies in the GOODS (Giavalisco et al. 2004) fields. We also introduced a statistical model which we used to simultaneously infer the spatial and

angular distribution of satellites along with their numbers. In this work, we expand our analysis to include the satellites of host galaxies selected from the COSMOS (Scoville et al. 2007) field which is approximately 20 times larger than the GOODS field although somewhat shallower. This dramatic increase in our sample size allows us to analyze the properties of the radial and angular distribution and the number of satellites, for satellites more than a thousand times fainter than their host galaxies. Furthermore, the greatly increased number of host-satellite systems allows us to analyze these properties in bins of redshift, host morphology, satellite luminosity and host mass. This is of key importance because, as has been discussed, the number and luminosity of satellites are believed to depend strongly on the properties of host galaxies. In order to make our results more easily relatable to theoretical predictions, we select our host samples by stellar mass and luminosity and analyze the number density of the satellite radial profile in units of R_{200} in addition to units of the half-width at half-max of the host light profile.

This Chapter is organized as follows: In §3.1, we describe the images and catalogs used in this analysis. In §3.2 we discuss the selection and properties of the host galaxy sample. In §3.3 we briefly review the host modeling and subtraction technique developed in Chapter 2 In §3.4 we present the binned radial and angular profiles of satellites to provide a qualitative sense of the signal. In §3.6 we discuss the theoretical model of the cumulative number of satellites per host as a function of host luminosity from Busha et al. (2011) which we compare with our measurements. In §3.7 we explain how we tested our statistical analysis and its limitations. In § 3.8 we present the results. In § 3.9 we compare our results with similar studies performed at low redshift. In § 3.10 we discuss the broader implications of our results. In § 3.11 we provide a concise summary.

The Appendix contains more detailed explanations of many of the methods used in this chapter.

3.1 Imaging and Catalogs

To achieve the depth and area required to study satellite properties and their evolution, we use imaging and photometric catalogs taken from the COSMOS survey [1]. In Chapter 2 we used the deeper GOODS survey to develop and test our method. However, the GOODS survey is not sufficiently wide to allow us to perform a binned statistical analysis of the satellite population as a function of host properties as we do in this chapter.

The COSMOS field has extensive spectroscopic and photometric ground-based coverage which we use to catalog the redshift, stellar mass and morphology of the host galaxies. Spectroscopic redshift measurements from Lilly et al. (2007) are available for roughly 1/7 of the host galaxies. When spectroscopic measurements are not available we use ground-based photometric redshift measurements from the catalog by Ilbert et al. (2009), which is the same catalog we use for all stellar mass measurements. Finally for morphological categorization, we use the Cassata et al. (2007) morphological catalog. The Cassata et al. (2007) catalog is created by a computer algorithm which is known to fail in certain cases. For instance, spiral galaxies with large bulges can be misclassified as ellipticals. Thus we visually confirmed all morphological classifications using COSMOS ACS imaging. This resulted in the re-categorization of $\sim 10\%$ of the morphologies of

[1] COSMOS photometric catalogs and surveys are available at http://irsa.ipac.caltech.edu/data/COSMOS/datasets.html

host galaxies.

3.2 Host Galaxy Selection

In selecting our host galaxy sample, we had two main goals. The first was to ensure that we selected hosts that maximized our ability to detect neighboring satellites, and the second was to select a host sample that could be easily related to predictions from simulations. In Chapter 2, we achieved the first goal by selecting massive early-type host galaxies which had relatively smooth light profiles that made it easy to find nearby companions. However, galaxy morphology is not well reproduced by current simulations (e.g. Hoyle et al. 2011), making it difficult to select matching host galaxies from simulations. In contrast, simulations have had success matching the stellar mass function of massive galaxies, to dark matter halos in simulations (Berlind & Weinberg 2002; Zheng et al. 2007; Conroy & Wechsler 2009), making a mass-selected host sample more appealing.

To compromise between the two goals, we selected relatively high stellar mass, $\log[M^*_{\rm h}/{\rm M}_\odot] > 10.5$ host galaxies, as these have well studied stellar masses, redshifts and luminosities, and a greater number of satellites visible down to a fixed apparent magnitude (e.g. Kravtsov 2010; Busha et al. 2011). We study hosts in a redshift range between $0.1 < z < 0.8$. This range allows us to study evolutionary properties while still guaranteeing that we can detect satellites with luminosity contrast from the hosts equivalent to that of the SMC relative to the Milky Way at all redshifts. By doing this, we expect to observe approximately one satellite per host. We exclude $z < 0.1$ galaxies, which are few and too extended in angular size to analyze in the same way as the more

distant sample.

We also required host galaxies to be relatively isolated in order to ensure that they themselves are not satellites of larger central galaxies. This is important for our analysis because we do not want to count objects associated with the larger host as satellites of a smaller galaxy. Using the stellar mass and redshift catalogs, we include only host galaxies which are not within the R_{200} of a neighbor that has more than its stellar mass and is at the same redshift within measurement uncertainties. We calculate R_{200} given stellar mass using Equation 3 from Dutton et al. (2010), which provides a by eye fit to the observed and inferred relationships between stellar and halo mass of galaxies as a function of morphology. For early-type galaxies the best fit function is:

$$y = 10^{2.0} \left(\frac{x}{10^{10.8}}\right)^{-0.15} \left[\frac{1}{2} + \frac{1}{2}\left(\frac{x}{10^{10.8}}\right)^{2}\right]^{0.5} \tag{3.1}$$

And for late-types:

$$y = 10^{1.6} \left(\frac{x}{10^{10.4}}\right)^{-0.5} \left[\frac{1}{2} + \frac{1}{2}\left(\frac{x}{10^{10.4}}\right)\right]^{0.5} \tag{3.2}$$

Where $y = \langle M_{200} \rangle / M^*$ and $x = M^*$.

Although there is uncertainty in the stellar mass estimate and additionally in R_{200}, we consider these values to be know with absolute precision for the purposes of our analysis. We tested the effect of uncertain stellar mass on our analysis by doubling masses before calculating R_{200} and found no significant impact on our inference result. The final sample has 1901 early-type and 1524 late-type galaxies. The distribution of stellar masses, redshift and absolute r-band magnitudes are shown in Figure 3.1.

As discussed in the Introduction, there is theoretical and observational evidence that satellite properties may depend strongly on the properties of host galaxies. In order to

study these trends, we divide host galaxies into bins of redshift, morphology and stellar mass when performing our analysis of the satellite population. We choose two bins of stellar mass with $10.5 < \log_{10} M_{\rm h}^{*} < 11.0$ and $11.0 < \log_{10} M_{\rm h}^{*} < 11.5$, and two bins in redshift with 0.1 <z< 0.4 and 0.4 <z< 0.8.

When we analyze the satellite population, we study bins of cumulative magnitude contrast from the host galaxy, such that all satellites have magnitudes brighter than $m_{\rm sat} - m_{\rm host} < \Delta m$. In each bin of Δm we only study host-satellite systems such that the host is at least Δm magnitudes brighter than the limiting survey magnitude. This means a different subset of host galaxies is used to study $\Delta m = 8.0$ satellites than $\Delta m = 2.0$ satellites, leading to some variation in host properties within a stellar mass, redshift and morphology bin. In Tables 3.2 and 3.3, we summarize the average stellar mass, redshift and luminosity of the host galaxies used in each Δm analysis, split into bins of redshift, stellar mass and host morphology. Host galaxies used to study the faintest satellites tend to be at lower redshift than the average in a fixed redshift bin.

3.3 Detection and photometry of close neighbors

In order to study the satellite population, we require an accurate catalog of object positions and magnitudes as close as possible to the host galaxies. Companions of bright galaxies are difficult to study because they are intrinsically faint and often obscured by the host galaxy light. This is a serious issue when attempting to measure the slope of the power law-radial profile of the satellite spatial distribution, as the innermost regions of the system provide the best constraint on the slope. Furthermore, close to the host, the

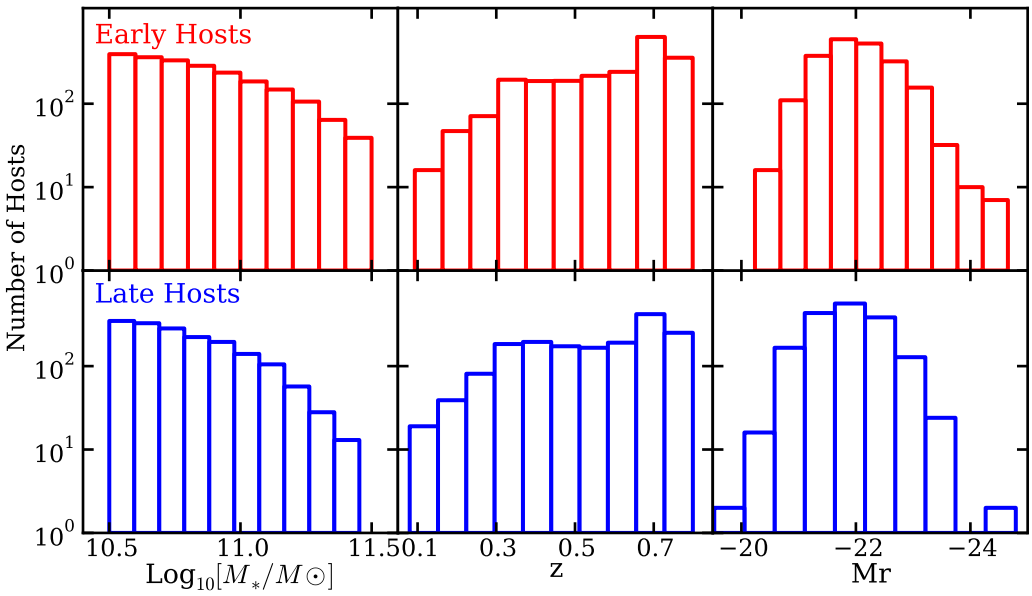

Figure 3.1: The distribution of stellar mass, redshift and absolute r band magnitude for *upper*: early and *lower*: late-type hosts.

ratio of the number of satellites to background/foreground galaxies is the most favorable. In Chapter 2, we developed a method of removing the smooth component of the host galaxy light profile to overcome some of these challenges. This process makes automated object detection much more accurate and reliable near the host galaxy. We use the results to update the COSMOS photometric catalog with newly detected objects, and to replace the photometry for objects that had already been detected near the hosts.

In this section we discuss the new objects this method allows us to identify in the COSMOS data. All SExtractor parameters used are listed in Table 3.4 in the Appendix.

3.3.1 Object detection and photometry

We detect objects in the host-subtracted images using SExtractor parameters tuned to match the object detection and photometry that was used to make the COSMOS catalogs. Appendix 3.12.1 contains a comparison of our photometry and object detection to that of the COSMOS catalogs in a large non host-subtracted field.

In order to test our sensitivity to low surface brightness objects, we simulated faint sources near our hosts with Sersic indices and effective radii and tested our recovery rate and photometric accuracy for these objects after host subtraction. Results of these simulations can be found in Figure 3.11 in the Appendix. We use the results from these simulations to identify a minimum radius at which we can detect at least 90 percent of simulated satellites with accurate photometry, and define this as the minimum radius at which we study the properties of the satellite population. We find that we can accurately recover satellites with MAG_AUTO $I_{814W} < 25.0$ as close as 2.5 R_h in COSMOS for the majority of hosts with early-type light profiles. This corresponds to a mean distance of $1\farcs2$ (7 kpc) with a standard deviation of $0\farcs7$ (3 kpc). The inner detection boundary is slightly higher for host galaxies with $M^*_{\rm h} > 11.0$ and $z < 0.4$ as these tend to have light profiles that extend above the background further from the host centers. Late-type galaxies have more extended light distributions in addition to spiral arms which are difficult to distinguish from neighboring galaxies, thus we choose a more conservative inner boundary of 4 $R_{\rm h}$(3$\pm$1 arcseconds, 17$\pm$ 6 kpc) for these hosts. On average, these minimum radii correspond to 0.02 and 0.07 R_{200} for early and late-type galaxies respectively.

3.3.2 Properties of Objects Detected in Cutout Regions

In this section we compare the properties of newly detected objects after host light subtraction to the properties of objects already in the COSMOS photometric catalog near the host galaxies. The upper panel of Figure 3.2 shows the distribution in the contrast in MAG_AUTO measurement (δm= $m-m_{\rm h}$)[2] between hosts and objects detected within 2.″5 of the host galaxies. The distributions are compared for objects already in the COSMOS catalog and newly detected objects. For both types of objects, photometry is performed after host light removal. Newly detected objects are about a magnitude fainter than objects that were already in the COSMOS photometric catalog, with average δm values of 3.22 compared to 4.44 for previously detected objects within the same region, with typical measurement uncertainty of 0.05 mag.

The lower panel of Figure 3.2 shows the number density of objects as a function of distance from the hosts. The number density of objects in the COSMOS photometric catalogs drops within the inner 1.″5, while the number density of newly detected objects rises, more than doubling the number of COSMOS detected objects in this region. The sum of the two number density signals increases steadily with decreasing distance from the host galaxy. Thus host subtraction and rigorous measurements of completeness in the innermost regions are necessary for an accurate measurement of the radial profile and number of objects close in projection to the host galaxies.

[2]Note the use of lower-case δm, which denotes a specific contrast from the host and is different from Δm which describes the allowed maximum contrast between host and neighboring objects for a particular data set.

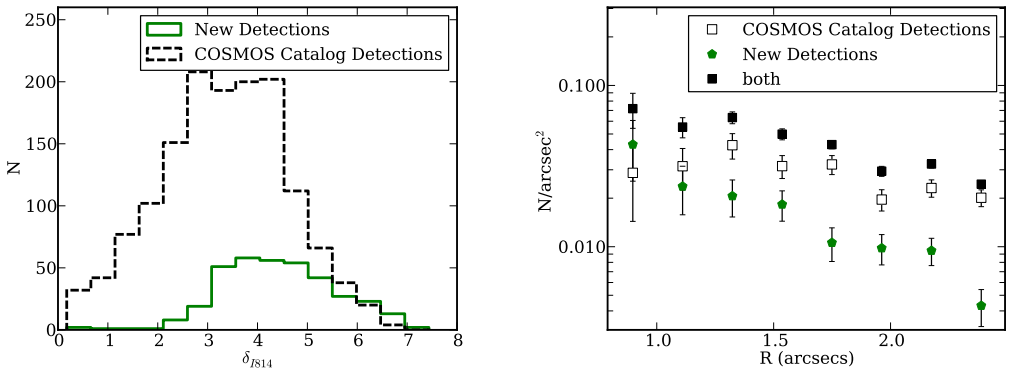

Figure 3.2: Comparison of the properties of objects detected in the COSMOS catalogs to those of newly detected objects after host light subtraction. *Left*: The distribution of magnitude differences from hosts (δm$= m - m_h$) within 2.″5. *Right*: The number density of objects as a function of distance from hosts. Newly detected objects are closer to the hosts than those in the COSMOS photometric catalog and make a significant contribution to the measurement of object number density within 2 arcseconds from the center of the host.

3.4 First Look

Before describing our model for the radial and angular profiles of objects near the host galaxies, it is instructive to show these distributions in spatial bins in order to provide a visual representation of the data. However, binning is inherently limited because it requires the averaging of data, thereby losing information. Furthermore it is not conducive to accurate subtraction of foreground/background galaxies over a range of redshifts. Thus we do not perform our analysis on the spatially binned data, but instead use this section to justify our model choices in Section 3.5.

3.4.1 Distance Scaling and Radial Distribution

We scale measured object distances to account for the range of redshift and host mass scales in our sample. A scale relating to the host light profile is the natural choice for the

observer as this will vary with host redshift as well as host mass according to the size-mass relation (e.g. Trujillo et al. 2006; Williams et al. 2010). For this distance scale, we use $R_{\rm h}$ which is AWIN_IMAGE from SExtractor . We also perform a parallel study with all distances scaled by R_{200} of the host galaxies. Unlike $R_{\rm h}$, R_{200} can be calculated in dark matter only simulations and is thus a better choice when attempting to compare results with simulations. However, estimating R_{200} requires multiband photometry and stellar mass modeling which is not always possible, so it is useful to perform this analysis using both distance scalings to see if one choice or the other leads to systematic differences.

Figure 3.3 shows the average number density of objects as a function of distance from the hosts, with distances scaled by $R_{\rm h}$ in the upper panel and R_{200} in the lower panel. The behavior is qualitatively similar for both choices of distance scaling; the number density of sources increases as a power-law near the hosts. At large radii, the number density becomes dominated by the isotropic and homogeneous distribution of objects not associated with the hosts, represented by the gray dashed lines.

In Section 3.5 we describe how we analyze the number density signal by inferring the combined properties of the satellite and background/foreground populations. In Section 3.8.1, there is a comparison of the results using the two distance scalings.

3.4.2 Angular Distribution

In Figure 3.4 we show the angular distribution of objects within 10 $R_{\rm h}$, plotted for all hosts, early-type hosts and late-type hosts, where $\theta = 0$ is aligned with to the major axes of the host light profiles. This figure only includes host galaxies with axis ratio

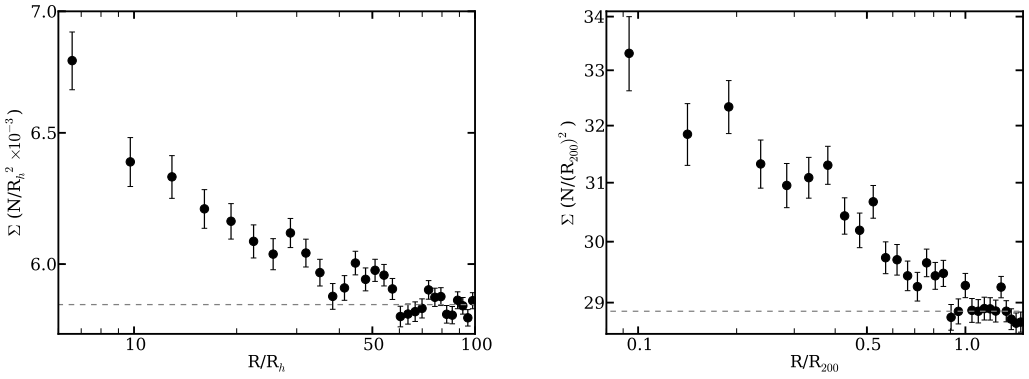

Figure 3.3: The average number density of objects near hosts as a function of radial distance. *Left*: In units of the second order moment of the intensity of the host-light profile along its major axis ($R_{\rm h}$), and *Right*: In units of R_{200} estimated from the host stellar mass.

b/a<0.6, to ensure that the direction of the host major axis is clearly measurable. As background/foreground objects are expected to be distributed isotropically relative to the host galaxy, any anisotropy we observe is caused by correlated structure presumably, in the form of satellites. In this region we expect a significant contribution to the number density to come from satellites, as evidenced by the strong satellite signal within this region in the upper panel of Figure 3.3.

Near early-type galaxies, the angular distribution shows a dominant component aligned with $\theta = 0$. A Komogorov-Smirnoff KS test gives a probability of $\sim 10^{-8}$ that the objects near early-type galaxies have a uniform angular distribution. In contrast, the objects near late-type galaxies appear more isotropically with a KS probability of being uniform of a few percent. This simple examination has been done without any effort to remove background/foreground contamination. However, it indicates that the satellites of early and late-type hosts may have different angular distributions relative to their host light profiles. To test this further, we separate satellite populations based on host morphology

in our inference of the parameters of the satellite spatial distribution.

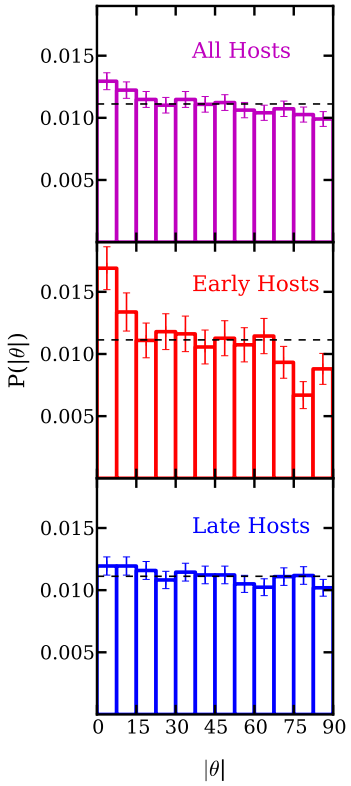

Figure 3.4: The average number of objects at angle $|\theta|$ from the major axis of the host light profile, within 10 $R_{\rm h}$, for hosts with elongation (b/a) less than 0.6. The distribution of objects near early-type hosts is more aligned with the host light profile than the distribution of objects near late-type hosts.

3.5 Joint Modeling of Satellite and Background Galaxy Populations

3.5.1 Background/Foreground Objects

We model the background/foreground number density signal around each host as a homogeneous, isotropic signal, with mean surface density Σ_b with magnitudes brighter

than the field magnitude limits ($I_{814W} < 25.0$). We inform our inference of Σ_b by measuring the local background around each of our hosts and using the mean and standard deviation of this distribution to create a prior.

We follow the method recommended by Chen et al. (2006) of estimating the local background around our host galaxies (rather than taking the field average) in order to accurately measure Σ_b. This is important for removing correlated line of sight structure which is not within the virial radii of the host galaxies. In Chapter 2, we measured the background between 45 and 70 $R_{\rm h}$, where no significant satellite signal was apparent. For consistency, in this analysis, we estimate the background in the same region when scaling distances by $R_{\rm h}$. When performing the analysis in units of R_{200}, we calculate the background between 1.0 and 1.5 R_{200}, following the simulation results from Liu et al. (2011).

Various studies have shown that the local projected number density of objects is correlated with host stellar mass and morphology. To account for this, we measure the mean and standard deviations separately for each data set that we run the inference on and use this local background information as a prior when inferring the properties of the satellite distribution.

For a given value of Δm and host galaxy magnitude, $m_{\rm h}$, only a fraction of the total number of background/foreground objects will appear in our analysis. To calculate this fraction, we model the cumulative distribution function (CDF) of the background number counts by a power-law (e.g. Benítez et al. 2004). Defining $\Sigma_{\rm b,o}$ as the background density measured for all objects with magnitudes brighter than the limiting survey magnitude $m_{\rm lim}$, the number density of background/foreground objects around the $j^{\rm th}$ host is given

Parameter	Description	Prior
Satellite Model		
Ns	Number of satellites per host	U(0,20)
$\gamma_{\rm p}$	Logarithmic slope of the satellite radial distribution	U(-10,0)
A	Relative flattening between satellite angular dist. and host light	U(0,2)
$\|\phi\|$	Angle between the major axis of satellite angular dist. and that of the host light.	U(0,π/2)
Background Model		
$\Sigma_{\rm b,o}$	Number density of all background objects with $I_{814W} < 25$	varies with host mass and Δm
α_b	Logarithmic slope of the background number counts	N(0.300, 0.005)/N(0.305,0.005)

Table 3.1: Summary of model parameters and priors. U(a,b) denotes a uniform distribution between a and b and N(μ,σ) denotes a normal distribution with mean μ and standard deviation σ.

by:

$$\Sigma_{\rm b,j} = \Sigma_{\rm b,o} 10^{\alpha_{\rm b}(m_{\rm h,j}+\Delta m - m_{\rm lim})}. \tag{3.3}$$

We measure the background slope to be 0.305 ± 0.005 near low mass early and late-type galaxies and 0.300 ± 0.005 near high mass host galaxies.

3.6 Theoretical Prediction for the Number of Satellites per Host

Our choice to model the number of satellites as being constant for all hosts within a fixed magnitude contrast from the host magnitude is a simplification of a more complicated picture. Dark matter only simulations predict that dark matter halos should host an approximately constant number of dark matter subhalos with a given *dark matter mass fraction* of the host halo, regardless of the host halo mass (Kravtsov 2010, and references therein). This scale-invariance relates in a non-trivial way to our chosen observable which is the number of satellites we expect to see within a fixed magnitude contrast from from host halo. This is due to the non-linearity of the stellar-mass to halo-

mass relationship for hosts with $M_{\rm h}^{*} > 10^{10} M\odot$ as can be seen, for example, in Behroozi et al. (2010).

In recent work, Busha et al. (2011) (hereafter B11) modeled the number of satellites per Δm m bin as a function of host luminosity using SubHalo Abundance Matching (SHAM) techniques (Kravtsov et al. 2004a; Behroozi et al. 2010), to connect dark-matter halos in the Bolshoi simulations, (Klypin et al. 2011; Trujillo-Gomez et al. 2011) to the r-band luminosity function of galaxies. In order to properly mimic the observational selection function in this work, we have reproduced the measurements of B11 using the luminosity function from the AGES simulations (Cool et al 2012) as applied to the Bolshoi halo catalog at the appropriate redshift.

To test how the inference method responded to the non-Poissonian distribution of satellites predicted by theory we ran simulations using the distribution of host luminosities in our sample to generate the number of satellites per host based on Equation 8 from B11. Spatial positions for satellites and background/foreground objects were drawn stochastically from the model described in Section 3.5. As desired, our inference accurately returned the mean number of satellites per host along with the other input parameters in the simulations.

3.7 Analysis

In Chapter 2, we provided details for the construction of the posterior probability distribution function (PDF) which allow us to infer values for the parameters of our model given the data and prior knowledge of the background listed in Table 3.1. For

each parameter in each subset of the data listed in Tables 3.2 and 3.3, we compute the posterior PDF using a Markov Chain Monte Carlo (MCMC) method. At least 10^4 iterations per chain are performed in order to ensure convergence.

To study variation with host properties, satellites are analyzed in bins of 'high' and 'low' host stellar mass corresponding to $10.5 < \log_{10}[M^*_{\rm h}/M\odot] < 11.0$, $11.0 < \log_{10}[M^*_{\rm h}/M\odot] < 11.5$, as well as 'low' and 'high' host redshift corresponding to $0.1 < z < 0.4$ and $0.4 < z < 0.8$ and early and late-type host galaxies. Every bin in Δm is analyzed separately for each bin in host morphology, redshift and stellar mass, using an appropriate prior on $\Sigma_{\rm b,o}$ estimated using the local background for that subset of host galaxies.

In order to combine results from different data sets, we bin the posterior PDFs for the parameters of interest and multiply them together. When data sets were analyzed in different regions around the host, we first re-normalize the inferred satellite numbers using the posterior median values of $\gamma_{\rm p}$ and $N_{\rm s}$ to account for the differences in examined areas, before combining the posteriors.

We tested our inference by running simulations with varying background and satellite properties. We find that in order to accurately infer the radial profile $\gamma_{\rm p}$ we need a minimum of 50 host galaxies *and* at least 20 satellites. The first requirement guarantees that the inference has an adequate estimate of the background density, and the second that there is a sufficient satellite signal. Inferring the angular distribution of satellites is more difficult as this requires two dimensions of information. We find that at least 50 satellites are necessary to recover A and $|\phi|$. Furthermore, the inference on $|\phi|$ becomes inaccurate for values of A less than 0.5, thus we do not report confidence intervals of $|\phi|$

where the posterior favors values of A less than 0.5.

3.8 Results

We divide this section into the three characteristics of the satellite population that our inference constrains, namely the radial and angular distribution of satellites as well as the cumulative luminosity function. For each characteristic we discuss variation with host morphology redshift and stellar mass as well as with satellite luminosity by comparing results in high and low redshift and stellar mass bins as well as in bins of host morphology, as discussed in Section 3.7. Summaries of the inference results for each subsample Tables 3.2 and 3.3 in the Appendix. To show typical results, Appendix 3.13.1 contains the full bivariate posterior PDFs for all model parameters for high mass, early and late-type hosts for the $\Delta m = 4.0$ bin.

3.8.1 Radial Distribution

In Figure 3.5 we show the inferred value of the projected slope of the radial profile of the satellite number density as a function of Δm, divided by host morphology, stellar mass and redshift. We further compare these results when distances are scaled by $R_{\rm h}$ and by R_{200}.

This figure contains several important results. The first is that there is a significant detection of a population of objects with a power-law radial distribution up to 6.5 magnitudes fainter than low redshift hosts, and 5.5 magnitudes fainter than high redshift hosts. Using the average luminosity of the host galaxies in each of those data sets, this

corresponds to approximate absolute r band magnitudes of -16.1 and -17.4, respectively. The former is about one magnitude brighter than Sagittarius, the latter is similar to the absolute magnitude of the present day SMC.

Second, there is no significant difference in the inference of $\gamma_{\rm p}$ when distances are scaled by $R_{\rm h}$ or by R_{200}. This is remarkable given the number of assumptions that go into estimating the virial mass given the stellar mass, and is a very useful result for comparisons between theoretical and observational work for future satellite surveys where accurate host stellar masses are not available. For simplicity, for the remainder of the chapter we discuss results only for distances scaled by R_{200}. Results for both distance scalings are listed in Tables are listed in Tables 3.2 and 3.3 in the Appendix.

Third, there is no significant variation in $\gamma_{\rm p}$ as Δm varies with fixed host properties over ranges as large as 5.5 magnitudes. This is consistent with the physical processes governing the satellite luminosity function being not spatially related to the host galaxy over this fairly large luminosity range within the distance scales we study, given our measurement uncertainties.

The average values of $\gamma_{\rm p}$ for host galaxies of all morphologies are: -1.1 ± 0.1, -1.3 ± 0.4, -1.2 ± 0.1 and -1.0 ± 0.1 for satellites of low mass-low redshift hosts, high mass-low redshift hosts, low mass-high redshift hosts and high mass-high redshift hosts respectively.

Individual subsets of host morphology do not show any significant deviation from these values.[3] As there are no significant differences in the results for these samples, we argue that the satellite population can be well described by a power law with $\gamma_{\rm p} = -1.1 \pm 0.3$

[3]Although the high mass low redshift sample of late-type galaxies appears to have a steeper satellite number density radial profile, this sample is relatively small (only 54 hosts) and does not show a statistically significant deviation.

within the level of precision afforded by our data.

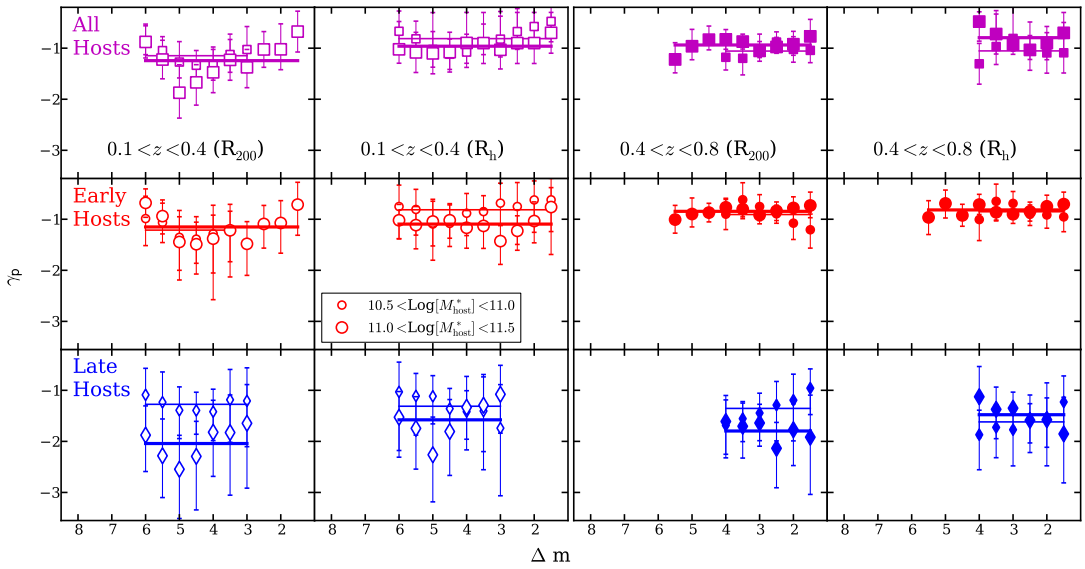

Figure 3.5: Projected radial profiles of the satellite number density distribution with distances scaled by $R_{\rm h}$ and R_{200} for satellites divided by host morphology, redshift and stellar mass. Horizontal lines indicate the average median value of $\gamma_{\rm p}$. Large markers indicate results for satellites of higher stellar mass hosts and small markers for the satellites of lower stellar mass hosts.

3.8.2 Angular Distribution

The satellites of early and late-type galaxies display markedly different angular distributions where the sample of satellites and host galaxies is large enough to allow an inference on the parameters on the angular distribution. As an example, in Figure 3.6 we show the two dimensional bivariate posterior distributions for the parameters A and $|\phi|$ for satellites with $\Delta m < 4$ for early and late-type hosts with $0.4 < z < 0.8$ and 11.0 <Log[M*]< 11.5. The results for this subsample which are representative of all non-uniform posterior PDFs. One sided confidence intervals for A and $|\phi|$ for other non-uniform subsamples are listed in Tables 3.2 and 3.3.

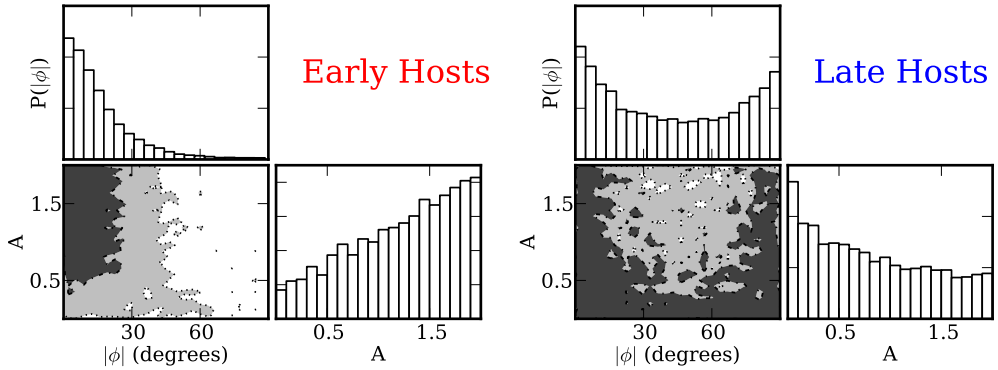

Figure 3.6: Bivariate posterior distributions of $|\phi|$ and A for the satellites of *Left:* early and *Right*: late-type host galaxies, for hosts with $0.4 < z < 0.8$ and $11.0 < \log[M^*_{\rm host}/M\odot] < 11.5$

Satellites of early-type galaxies show strong anisotropy, with the most likely value of A at 2 and $A > 0.9$ at a 68% confidence level. This indicates that the angular distribution of satellites is more flattened than the light profiles of the early-type galaxies. Furthermore, the satellites of early-type galaxies, are found preferentially along the major axis of the host light profiles. The offset, $|\phi|$ has a most likely value of 0 and is less than 20 degrees at a 68% confidence level.

In contrast, the average angular distribution of the satellites of late-type galaxies is consistent with being isotropic relative to the host disks, with A having a most likely value of 0 and $A < 1.1$ at a 68% confidence level. As discussed in Section 3.7, the inference on $|\phi|$ is unreliable for small values of A as $|\phi|$ has decreasing influence on the likelihood function for small values of A.

There is a concern with regards to describing the ellipticity of the angular distribution of satellites as a fraction, A, of the host ellipticity (see Equation 2.5). Namely, late-type galaxies are much more elongated than early-type galaxies but there is no reason to

expect the satellites to reflect this. It may be that the satellites of late-type galaxies have anisotropic distributions on average but that the flattening of their distribution is less than the extreme flattening of the disks, making them look isotropic in comparison. To test whether the inferred isotropy is due to the extreme flattening of late-type hosts, we re-ran the inference on late-type hosts, this time artificially changing any late-type axis ratio b/a to 0.6 that was previously flatter than 0.6. We found that the posterior PDF still favored $A = 0$. However, the inference on the angular distribution for late-type hosts is difficult due to the large region obscured by spiral arms. To try to maximize the signal to noise, we performed a combined inference on late-type hosts of all redshifts and stellar masses. In this case, there was evidence for slight anisotropy, with satellites aligned with the galactic disk. Thus, an analysis of the satellites of late-type hosts with a large data set is warranted before any strong conclusions are reached.

3.8.3 Cumulative Luminosity Function

We detect a significant population of satellites as faint as 6.5 magnitudes fainter than their host galaxies. Results for the inferred number of satellites per host as a function of maximum contrast between host and satellite magnitude ($m_{\rm host} - m_{\rm sat} < \Delta_m$) can be found in Tables 3.2 and 3.3. A dominant uncertainty in this analysis is caused by the covariance between $\gamma_{\rm p}$ and $N_{\rm s}$ as can be seen in Figure 3.12; the same total number of objects can be achieved if there are more satellites following a shallower radial profile or if there are fewer satellites with a very steep radial profile, and more background objects. We alleviate this degeneracy by performing the inference a second time, using a Gaussian

prior on $\gamma_{\rm p}$ from the results in Section 3.8.1. Given that $\gamma_{\rm p}$ showed no significant variation given measurement uncertainties as a function of satellite magnitude, host redshift or host morphology (see Section 3.8.1) we apply the same Gaussian prior on $\gamma_{\rm p}$ to all data sets, with mean -1.1 and standard deviation of 0.3.

Results for the inferred number of satellites using the prior are consistent with results without the prior and are listed in Tables 3.2 and 3.3 The prior on γ_p allows for a significant detection of satellites up to 8 magnitudes, or more than a thousand times fainter than their host galaxies for the low reshift host sample. Using the average absolute r band magnitude of hosts, this corresponds to satellites with $M_r \sim -13.5/-14.7$ for low and high mass host samples respectively, and is similar to the absolute magnitude of Fornax. The range of Δm does not change for the higher redshift sample but the average measurement uncertainty is decreased.

In Figure 3.7 we plot the number of satellites per host, $N_{\rm s}$, in increasing bins of Δm, for varying host morphologies, stellar masses and redshifts between 0.07 and 1.0 R_{200} ($\sim 17-200$ kpc). The points for 'All' morphology hosts come from binning and multiplying the posterior PDFs of the early and late-type hosts [4]. We always combine the posteriors even when the inference on the late-type host galaxies is not sufficiently determined to plot (i.e. for $\Delta m > 4$ for more massive, high redshift late-type hosts). The satellite CLF is a fairly constant power-law with slope approximately $N(L) \propto L^{-0.5\pm0.1}$. There is a slight upward shift in the normalization for satellites with $\Delta m > 6.5$, which is attributable in part to the slight increase in average host stellar masses for these bins

[4]This is valid as these data sets are independent. The product of the two posterior PDFs can be viewed as a measurement of the number of satellites per host for the entire sample

(see Tables 3.2 and 3.3). We leave a more detailed unbinned analysis of the satellite luminosity function to a future chapter.

There is a strong dependence between early-type host stellar mass and number of satellites within a fixed value of Δm. The dependence on stellar mass is less apparent for late-type host galaxies, with the more massive hosts having barely more satellites on average than less massive host galaxies despite the fact that the low and high mass late-type hosts have the same average stellar masses as the low and high mass early-type hosts. While there is a large difference between the number of satellites of the more massive early and late-type hosts, the numbers are the same for less massive hosts.

Figure 3.7 also shows the B11 theoretical prediction for the mean number of satellites given the distribution of host luminosities in each sub-sample. The dotted line represents an extrapolation of the theoretical model, to luminosities where the simulation used begins to suffer from incompleteness effects. The theoretical model is in excellent agreement with observation for objects with $0.1 < z < 0.4$, while it tends to underpredict the abundance of satellites of higher redshift objects (although the slopes remains in good agreement). It should be noted, however, that the luminosity function used in the model in the range $0.4 < z < 0.8$ suffers from significant observational uncertainties. Such uncertainties directly impact halo occupation in non-trivial ways, since parameters such as M_* and ϕ_* have a significant impart on the mass-luminosity relation for massive objects. Additionally, we assumed a fixed scatter in the mass-luminosity relation of 0.16 dex, which has been shown to be in good agreement with z = 0.1 galaxies (Behroozi et al., 2010), but has yet to be explored at higher redshift. A full exploration of the theoretical uncertainties relating to this prediction is beyond the scope of the current chapter and

is left for future studies.

As expected, subhalo abundance matching predictions which are constructed without taking host morphology into account cannot capture the morphological dependencies. A possible explanation for the variations is given by the morphology-density relation (Dressler 1980; Postman & Geller 1984; Treu et al. 2003), in the sense that at fixed luminosity, early-types tend to reside in denser environments and have more massive dark matter halos than late-types. This in turn might affect the properties of luminous satellites and points to limitations of the present SHAM models.

3.9 Comparison with previous work

We can make several comparisons between our work and low-redshift studies of the satellite population. To start, our inference on $\gamma_{\rm p} = -1.1 \pm 0.3$ is consistent with results from the low redshift SDSS study of LRGs by Watson et al. (2010) and of satellites with luminosities similar to that of the LMC (Tollerud et al. 2011). It is slightly shallower, even though marginally consistent, than what was measured by Chen (2008). Given uncertainties in the scale radii of host galaxies, we are unable to determine whether our result is consistent with NFW as was observed by Guo et al. (2012). Our result is consistent with our previous work (Chapter 2), in which we studied the satellites of early-type galaxies from the GOODS fields.

In contrast to Watson et al. (2012), we do not find a significant trend between Δm and $\gamma_{\rm p}$ over the 5 magnitudes in Δm that we studied. While Watson et al. (2012) measured the satellite radial profile changing from $\gamma_{\rm p} \sim -2$ for satellites with $M_r = -21$

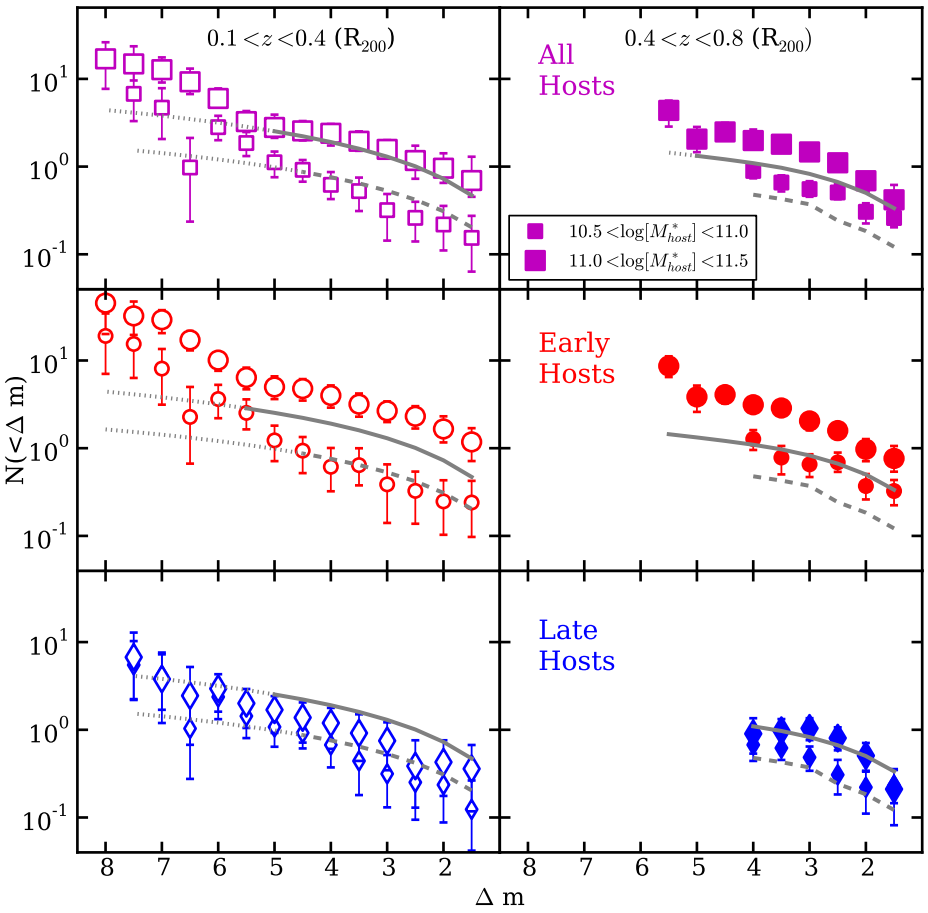

Figure 3.7: The cumulative number of satellites per host between 0.07 and 1 R_{200} as a function of the magnitude contrast between host and satellite galaxies, plotted for different samples of host redshift, stellar mass and morphology. Purple squares, red circles and blue diamonds represent all, early and late-type hosts respectively. The gray solid and dashed lines are theoretical predictions for the satellites of high and low mass host galaxy samples respectively. Thin dotted lines indicate an extrapolation of the theoretical prediction which was made for satellites brighter than $M_r < -17$. Note that the mean host stellar mass and redshift within each bin shifts slightly towards higher masses and lower redshifts starting at $\Delta m = 6$ (see Tables 3.2 and 3.3). .

to $\gamma_{\rm p} \sim -1$ for satellites with $M_r = -18.5$. The closest matching data-set in this study is the low-redshift, high-mass sample, for which we infer $\gamma_{\rm p} = -0.7^{+0.4}_{-0.6}$ for $\Delta m = 1.5$ satellites (which is approximately $M_r \sim -21$) and $-1.7^{+0.6}_{-0.9}$ for $\Delta m = 4.5$ satellites (approximately $M_r \sim -18.5$). The difference may be due to the fact that our enhanced

detection technique allows us to detect a significant new population of faint satellites which would have otherwise remained obscured, causing the radial profile to appear to be flatter. Given our measurement uncertainties we cannot rule out a small increase in the concentration of fainter satellites towards the center as was observed by Tal et al. (2011) and Guo et al. (2012).

The dependence of the satellite angular distribution on host morphology is consistent with numerous previous low-redshift studies of bright satellites (Bailin et al. 2008; Brainerd 2005; Agustsson & Brainerd 2010), which also found that satellites of early-type hosts tend to be found along the major axis of the host light profiles. Furthermore, these works found no significant anisotropy in the angular distribution for the satellites of late-type hosts (see also Yegorova et al. 2011).

The satellite numbers we measure for lower redshift hosts are consistent with the results from low redshift SDSS studies of hosts with similar masses. In Figure 3.8, we compare the number of satellites of low mass, low redshift host galaxies to the number in the Milky Way and near other low mass, low redshift host galaxies by Guo et al. (2011a), Liu et al. (2011), and Strigari & Wechsler (2012). The Milky Way has an absolute magnitude of $M_r \sim -21.2$, which is slightly lower than the typical host luminosity in the low stellar mass, low redshift data set in this sample ($M_r \sim -21.6$). However we expect some passive evolution in the stellar mass to light ratio of order one between redshifts 0.8 and 0.1 (e.g. Treu et al. 2005b), which would make the average luminosity approximately -21.3 at present day. Furthermore, the typical stellar mass in the low mass, low redshift subset of galaxies is $\log_{10}[M_h^*/M\odot] \sim 10.7$ which is equivalent to the stellar mass of the Milky Way within measurement uncertainties (McMillan 2011). Thus the hosts are

approximate Milky Way analogs at redshift 0.3.

In Figure 3.8, we also plot two theoretical predictions. The B11 subhalo abundance matching is identical to that plotted in the corresponding panel in Figure 3.7. The second shows the 10-90 percent tails of the predicted satellite distribution from Guo et al. (2011b) semi-analytic models of galaxy formation applied to the Millennium and Millennium II simulations (Springel et al. 2005; Boylan-Kolchin et al. 2009).

At all comparable magnitudes our inferred number of satellites is consistent with that measured by Guo et al. (2011a), Liu et al. (2011) and Strigari & Wechsler (2012). At the bright end, there are on average fewer satellites per host than what is observed in the Milky Way. The infrequent presence of Magellanic Cloud-equivalent satellites ($\Delta m = 2.0$) of Milky Way mass hosts has been noted in numerous surveys before this (Guo et al. 2011a; Lares et al. 2011; Liu et al. 2011). We are able to measure the satellite number down to 8 magnitudes fainter than the host galaxies and corresponds to an average r band luminosity of about -13.3 which is similar to the Fornax satellite. At the faint magnitudes the satellite numbers fall just under the upper limits from Strigari & Wechsler (2012) and the slope of the luminosity function is consistent with that at brighter satellite magnitudes.

At higher redshifts, Newman et al. (2012) measured the pair fraction of galaxies $\Delta m < 2.5$ between redshifts 0.4 and 2. Scaled to the same region as our study, they found $N_{\rm s} = 0.48 \pm 0.09$. This number is somewhat higher than our measurement for low-mass host galaxies which had 0.14 ± 0.1 in the same region. This discrepancy is likely due to the fact that the authors used global rather than local background subtraction, meaning they estimated the average background across the entire field. This leads to

a lower estimate of the number density of background objects than local background estimation as it does not take into account the filamentary clustering of galaxies. The authors estimate that using local rather than global background estimation increases their pair fraction by a factor of two relative to local background techniques, which brings the results into agreement (0.24 ± 0.05 compared to our measurement of 0.14 ± 0.1). The authors observed no significant evolution in the pair fraction in the redshift interval they studied.

In Figure 3.9 we show the satellite CLF for varying host luminosities and redshifts. For comparison, we include the measurement of intermediate luminosity hosts from Guo et al. (2011a) at redshift 0.1 along with the theoretical prediction from B11 for redshift 0.3 hosts . This figure highlights the the strong dependence of the number of satellites per host on the host stellar mass and luminosity. At the same time, there is no significant redshift evolution in the number of satellites. This highlights the important of selecting constant host mass samples when attempting to study trends in the satellite population.

As shown by the curves, theoretical predictions using the techniques of B11 match our inferred satellite numbers well for the sample of 'All' morphology hosts at $z = 0.3$, as discussed above.

3.10 Discussion

In the Introduction we discussed a large number of physical processes that can be tested by the spatial distribution and number of satellite galaxies. These processes can broadly be divided into three sets. First how satellites are affected by interactions with

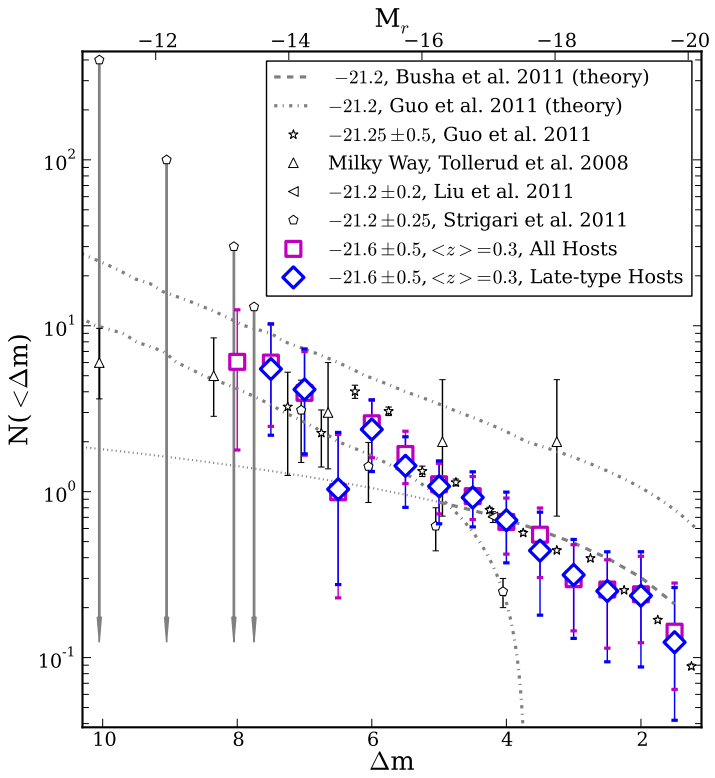

Figure 3.8: Summary of measurements of the satellite population of Milky Way-luminosity hosts. The legend lists the mean value and variation in the r band absolute magnitude for each study. The theoretical prediction from Guo et al. (2011b) represents the 10 and 90 percentiles of the satellite probability distribution. Low-mass, low-redshift hosts in this work have a mean redshift of 0.3 and $< M_r >\sim -21.6$ which is brighter than the present day Milky Way. However, accounting for passive evolution of roughly one magnitude from redshift one to present, these galaxies represent approximate Milky Way analogs at redshift 0.3. The upper x-axis is shifted to the left by 0.4 mags for these hosts.

their host galaxies, second how host galaxies are affected by satellites, and third how our observations fit into the framework of ΛCDM.

In terms of the first question, we see no evidence for strong physical interaction between satellite galaxies and central host galaxies on the distance scales we study ($\sim 10 - 250$ kpc). This is evidenced by the lack of any kind of trend in the steepness of the radial profile of the satellite number density as a function of host redshift, stellar mass,

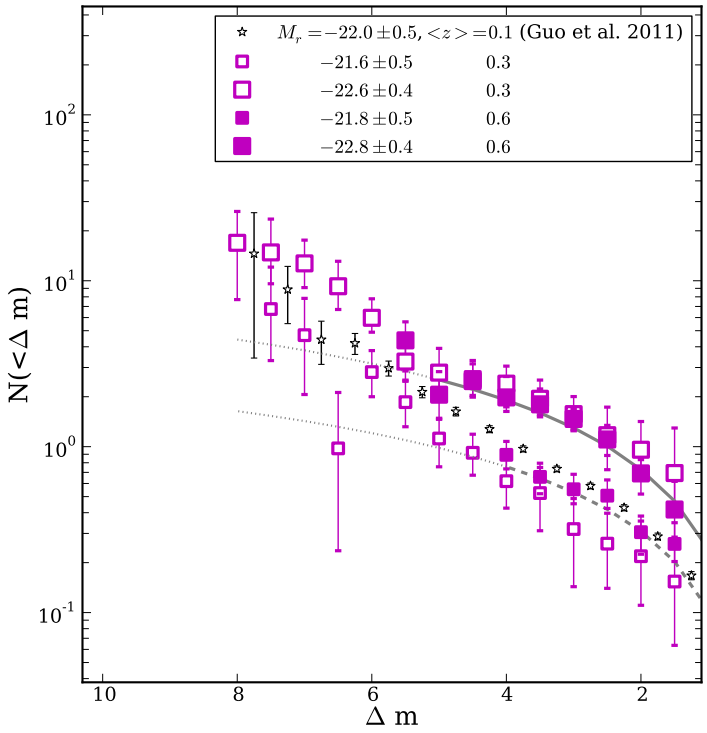

Figure 3.9: Comparison of the cumulative luminosity function of satellites of hosts with varying luminosities. Mean host absolute magnitudes and redshifts are listed in the legend. The gray dashed and solid lines are theoretical predictions for high and low mass, low redshift host galaxies from this work. The satellite numbers depend strongly on host luminosity, while there is no significant trend with redshift given measurement uncertainties.

or morphology or as a function of satellite luminosity. In a future paper, we will allow the slope of the radial profile to vary with distance from the host galaxies to allow us to compare the profile with NFW. We see evidence that the satellite angular distribution follows the host mass profile. Assuming that the satellite number density follows the dark matter halo, our result indicates that early-type host galaxies have light profiles that are aligned with their dark matter halos. The alignment for satellites of disk galaxies is weaker indicating that disk galaxies may be less aligned with their dark matter halos.

This interpretation of morphologically dependent galaxy-halo alignment is consistent with results from weak lensing (Mandelbaum et al. 2006), and recent theoretical modeling of disk galaxy formation (Deason et al. 2011). The fact that satellites of early-type hosts appear in elliptical distributions which are more flattened than the host galaxy light profile may indicate that the extended dark matter halo is more elongated than the galaxy light profile at the center. Several simulations show that baryons can have a dissipational effect on the central dark matter halo, causing it to be more isotropic than it would have been in absence of baryons (Dubinski 1994; Kazantzidis et al. 2010).

With respect to the second question, the constant radial distribution and number of satellites over time indicate a constant accretion of satellites for all host galaxies between redshifts 0.8 and 0.1 for satellites brighter than $\Delta m < 5.5$. Newman et al. (2012) observed no evolution for bright $\Delta m < 2.4$ pairs between 0.4 and 2.0. The combination of these results indicate that the number of satellites per host has been in equilibrium for at least half the age of the universe. From this we expect that host galaxies have been accreting stellar mass at a fairly constant rate over this time. Newman et al. (2012) estimate that the added stellar mass from the bright pairs is roughly $\sim 6\%$ per Gyr.

The strong dependence on host mass in the normalization of the satellites of early-type galaxies within a fixed bin of Δm reflects the non-linear relationship between host stellar mass and halo mass. The fact that late-type hosts do not show a significant corresponding trend may indicate the effects of environment. As discussed above, from the morphology-density relation, early-type hosts are more likely to be found in groups, while late-type hosts exist in more isolated environments. This suggests that dark matter halo mass-luminosity relation may also be dependent on environment, something that has

been largely ignored in most theoretical models. The relationship between host mass and satellite numbers is fundamental to the missing satellite problem, as clusters of galaxies have satellite numbers similar to what is predicted by ΛCDM while isolated field galaxies do not (Kravtsov 2010; Klypin et al. 1999).

Gravitational lensing is currently the only viable method of measuring the halo mass of satellite galaxies outside of the Local Group. Dalal & Kochanek (2002) used flux ratio anomalies in five quadruply-lensed quasars to estimate the fraction of dark matter in satellite galaxies to be between 0.6 and 7 % near the lensed images, demonstrating the potential of this technique to test the smallest mass scales of ΛCDM at cosmological distances. More recently, Vegetti et al. (2010a, 2012) used reconstructions of the mass profile of lens galaxies to detect low mass sub-halos near the lensing galaxies. From this Vegetti et al. (2012) estimated a mass fraction of $3.3^{+3.6}_{-1.8}$ in satellites near lensed images. Simulations find a somewhat lower mass fraction in satellites near the Einstein radius (Xu et al. 2010) of 0.1%.

We can use the results from this work to make an estimate of the fraction of satellites we would expect near massive, intermediate redshift, early-type galaxies (typical of lenses). Assuming a typical Einstein radius of $\sim$ 1", we expect roughly 5 % of the more massive, early-type host galaxies in our sample to host a satellite with $\Delta m > 7.5$, indicating an average mass fraction in satellite galaxies of a few percent in this region, on average. As discussed in the Introduction, the most massive sub-halos are stripped of dark matter in the inner regions of the host halo. The discrepancy between simulation results and observation may therefore indicate the importance of baryons in the preservation of the most massive satellites close to the central galaxy.

Finally, the detection of significant anisotropy in the satellite angular distribution has important implications for using flux-ratio anomalies to detect satellites. The anisotropy increases the line of sight mass between a lensed image and the host galaxy, effectively making the host galaxy more massive in a particular direction by up to a factor of 6 (Zentner 2006). In a future paper, we will perform a detailed study of the predicted flux ratio anomalies caused by luminous satellites using the results from this work.

3.11 Summary

We employ the host light subtraction method developed in Chapter 2 to study the satellites of 3425 host galaxies selected from the COSMOS field. The depth of the ACS images allows us to measure the satellite luminosity function more than one thousand times fainter than host galaxies at low redshift, while the combination of host light subtraction and high resolution images allows us to accurately detect faint sources as close as 0.3 (1.4) arcseconds (kpc) to host galaxies. Using the large volume of the COSMOS field, we examine trends in the satellite radial distribution, angular distribution and luminosity function. We further examine how these trends vary with redshift, host stellar mass and host morphology. Our main results can be summarized as follows:

1. The number density of satellites of all host galaxies in the sample is well-described by a power-law $P(R) \propto R^{\gamma_{\rm p}}$ with $\gamma_{\rm p} = -1.1^{+0.3}_{-0.3}$. There is no significant deviation from this as a function of host morphology, stellar mass or redshift. Furthermore, there is no evidence for variation in $\gamma_{\rm p}$ with satellite luminosity.

2. The inference results are the same when satellite distances are scaled by the half-

width at half-max of the host galaxy light profile ($R_{\rm h}$) or by R_{200} determined from the host stellar masses.

3. Satellites of early-type hosts, follow angular distributions which show strong alignment with the major axis of the host light profile, and fall in an ellipse that is more flattened than the host light profiles. In contrast, the satellites of late-type hosts do not show strong anisotropy.

4. The satellite cumulative luminosity function (CLF) is a power-law with faint end slope approximately $N(L) \propto L^{-0.5\pm0.1}$.

5. The satellite CLF of early-type galaxies shows a strong mass dependence in the normalization, with more massive early-type hosts having significantly more satellites than less massive early-type hosts. A similar trend is not apparent for late-type host galaxies. This likely reflects the fact that massive early-type hosts are more likely to be found in groups and exist in more massive dark matter halos than their late-type counterparts. This highlights the mass dependence of the satellite luminosity function which is fundamental to the missing satellite problem.

6. There is no significant evolution evident in the CLFs of satellites between redshift bins of median 0.6 to median 0.3. Satellite numbers from the redshift interval 0.1-0.4 are consistent with lower redshift studies ($z < 0.1$) of the satellites of host galaxies with equivalent masses.

7. Predictions from subhalo abundance matching (Busha et al. 2011) broadly agree with the measured satellite numbers when host morphology is not considered. How-

ever, the theoretical predictions were not created to match trends with host morphology. We find that such models systematically overpredict the abundance of satellites surrounding massive late-type hosts, and underpredict the abundance around early-type hosts. As the observations show smaller differences for the satellite abundance around lower mass hosts ($10.5 < \log[M_h^*/M\odot] < 11.0$), the theoretical model tends to have smaller discrepancies. Additionally, agreement is better at lower redshifts ($z \sim 0.3$) then for higher redshifts ($z \sim 0.6$). As discussed, this is likely due to errors in the model predictions, due to uncertainties in the luminosity function and scatter in the mass-luminosity relation at these higher redshifts.

Acknowledgements We thank R. Wechsler, P. Schneider, D.D. Xu, L.V.E. Koopmans, G. Dobler and J. Ostriker for many insightful comments and suggestions. We thank O. Ilbert for providing a stellar mass catalog of objects in the COSMOS field, which greatly enhanced this work. We thank P. Capak and the rest of the COSMOS team for their work on the COSMOS ACS and ground-based images and catalogs. AMN and TT acknowledge support by the NSF through CAREER award NSF-0642621, and by the Packard Foundation through a Packard Fellowship. PJM was given support by the Kavli Foundations and the Royal Society in the form of research fellowships.

3.12 Appendix

Tables 3.2 and 3.3 list the one sigma confidence intervals for the inferred parameter values in the case of early and late type host galaxies respectively.

Δm	N_h	$\log[M_*/M_\odot]$ (a)	z(b)	M_r (a)	N_s (c)	γ_p (c)	A (c,d)	$\lvert\phi\rvert$ (c,d)	N_s (e)	γ_p (e)	Ns (c,f)
$10.5 < \log[M_*/M_\odot] < 11.0$, $0.1 < z < 0.4$, $0.02 < R/R_{200} < 0.5$, $2.5 < R/R_h < 45$											
1.5	210	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	...(g)	...	...	...	...	...	$0.13^{+0.1}_{-0.08}$
2.0	210	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	...	...	...	...	...	...	$0.14^{+0.1}_{-0.08}$
2.5	210	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	...	...	...	...	...	...	$0.2^{+0.1}_{-0.1}$
3.0	210	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	...	...	...	...	...	...	$0.2^{+0.2}_{-0.1}$
3.5	210	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	$0.3^{+0.2}_{-0.2}$	$-1.2^{+0.5}_{-0.9}$	...	...	$0.6^{+0.2}_{-0.2}$	$-0.9^{+0.3}_{-0.3}$	$0.4^{+0.2}_{-0.2}$
4.0	210	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	$0.2^{+0.3}_{-0.2}$	$-1.3^{+0.6}_{-1}$	> 0.7	< 48	$0.5^{+0.2}_{-0.2}$	$-0.9^{+0.4}_{-0.4}$	$0.3^{+0.2}_{-0.2}$
4.5	208	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	$0.4^{+0.3}_{-0.2}$	$-1.4^{+0.4}_{-0.5}$	> 0.7	< 24	$0.6^{+0.3}_{-0.3}$	$-1.1^{+0.4}_{-0.5}$	$0.5^{+0.2}_{-0.2}$
5.0	182	10.7 ± 0.1	0.31 ± 0.07	-21.7 ± 0.4	$0.5^{+0.4}_{-0.3}$	$-1.4^{+0.5}_{-0.7}$	> 0.9	< 26	$0.7^{+0.4}_{-0.4}$	$-1.1^{+0.5}_{-0.7}$	$0.7^{+0.3}_{-0.3}$
5.5	182	10.7 ± 0.1	0.31 ± 0.07	-21.8 ± 0.4	$1.4^{+0.6}_{-0.6}$	$-1.1^{+0.4}_{-0.4}$	> 0.7	< 31	$1.6^{+0.8}_{-0.7}$	$-0.8^{+0.4}_{-0.5}$	$1.4^{+0.6}_{-0.5}$
6.0	58	10.8 ± 0.1	0.23 ± 0.06	-21.8 ± 0.4	2^{+1}_{-1}	$-1.0^{+0.4}_{-0.5}$	> 0.7	< 51	2^{+1}_{-1}	$-0.8^{+0.4}_{-0.6}$	$2.0^{+0.9}_{-0.6}$
6.5	28	10.8 ± 0.1	0.19 ± 0.05	-21.6 ± 0.4	...	...	...	...	...	...	$1.3^{+2}_{-0.9}$
7.0	18	10.9 ± 0.1	0.18 ± 0.05	-21.7 ± 0.5	...	...	...	...	...	...	5^{+3}_{-3}
7.5	5	10.9 ± 0.1	0.13 ± 0.03	-21.5 ± 0.4	...	...	...	...	...	...	9^{+7}_{-5}
8.0	3	10.9 ± 0.1	0.12 ± 0.01	-21.5 ± 0.3	...	...	...	...	...	...	11^{+9}_{-7}
$11.0 < \log[M_*/M_\odot] < 11.5$, $0.1 < z < 0.4$, $0.05 < R/R_{200} < 0.5$, $4 < R/R_h < 45$											
1.5	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$0.6^{+0.3}_{-0.3}$	$-0.7^{+0.4}_{-0.6}$	> 0.7	< 47	$0.5^{+0.3}_{-0.3}$	$-0.8^{+0.4}_{-0.5}$	$0.6^{+0.3}_{-0.2}$
2.0	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$0.8^{+0.3}_{-0.3}$	$-1.1^{+0.4}_{-0.6}$	> 0.7	< 36	$0.6^{+0.3}_{-0.3}$	$-1.0^{+0.4}_{-0.4}$	$0.9^{+0.3}_{-0.3}$
2.5	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$1.2^{+0.4}_{-0.4}$	$-1.1^{+0.4}_{-0.5}$	...	...	$0.9^{+0.3}_{-0.3}$	$-1.2^{+0.3}_{-0.4}$	$1.2^{+0.4}_{-0.3}$
3.0	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$1.1^{+0.5}_{-0.4}$	$-1.5^{+0.4}_{-0.6}$	...	...	$0.9^{+0.4}_{-0.4}$	$-1.4^{+0.3}_{-0.5}$	$1.6^{+0.5}_{-0.5}$
3.5	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$1.5^{+0.7}_{-0.7}$	$-1.2^{+0.4}_{-0.6}$	> 0.7	< 31	$1.2^{+0.6}_{-0.6}$	$-1.1^{+0.3}_{-0.4}$	$2.1^{+0.6}_{-0.6}$
4.0	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$1.7^{+0.8}_{-0.7}$	$-1.4^{+0.5}_{-0.7}$	> 0.9	< 20	$1.5^{+0.7}_{-0.7}$	$-1.2^{+0.3}_{-0.4}$	$2.1^{+0.6}_{-0.6}$
4.5	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$1.9^{+0.9}_{-0.7}$	$-1.5^{+0.4}_{-0.6}$	> 1.1	< 18	$1.7^{+0.7}_{-0.7}$	$-1.0^{+0.3}_{-0.4}$	$2.5^{+0.7}_{-0.7}$
5.0	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	$2.0^{+1.0}_{-0.9}$	$-1.4^{+0.5}_{-0.7}$	> 1.0	< 25	$1.4^{+0.8}_{-0.7}$	$-1.1^{+0.4}_{-0.5}$	$2.6^{+0.8}_{-0.7}$
5.5	92	11.2±0.1	0.32±0.06	-22.6 ± 0.4	3^{+1}_{-1}	$-0.9^{+0.3}_{-0.4}$	> 0.7	< 66	$1.6^{+0.9}_{-0.8}$	$-1.1^{+0.4}_{-0.5}$	$3.3^{+1.0}_{-0.9}$
6.0	81	11.2±0.1	0.31±0.06	-22.6 ± 0.4	6^{+1}_{-1}	$-0.7^{+0.3}_{-0.3}$	> 0.7	< 67	3^{+1}_{-1}	$-1.0^{+0.3}_{-0.4}$	5^{+1}_{-1}
6.5	39	11.3±0.1	0.27±0.07	-22.6 ± 0.4	9^{+2}_{-2}	$-0.7^{+0.3}_{-0.3}$	> 0.7	< 58	3^{+2}_{-2}	$-1.7^{+0.5}_{-0.9}$	9^{+2}_{-2}
7.0	15	11.3±0.1	0.21±0.04	-22.6 ± 0.4	...	...	...	...	...	...	15^{+4}_{-4}
7.5	9	11.3±0.1	0.19 ±0.04	-22.6 ± 0.4	...	...	...	...	...	...	17^{+8}_{-7}
8.0	3	11.3±0.1	0.19 ±0.04	-22.7 ± 0.4	...	...	...	...	...	...	23^{+16}_{-13}
$10.5 < \log[M_*/M_\odot] < 11.0$, $0.4 < z < 0.8$, $0.02 < R/R_{200} < 0.5$, $2.5 < R/R_h < 45$											
1.5	1038	10.7±0.1	0.6±0.1	-21.8 ± 0.5	$0.16^{+0.07}_{-0.06}$	$-1.2^{+0.3}_{-0.4}$	> 0.7	< 27	$0.17^{+0.06}_{-0.05}$	$-1.0^{+0.3}_{-0.4}$	$0.18^{+0.06}_{-0.06}$
2.0	1019	10.7±0.1	0.6±0.1	-21.8 ± 0.5	$0.21^{+0.08}_{-0.08}$	$-1.1^{+0.3}_{-0.3}$	...	...	$0.20^{+0.05}_{-0.06}$	$-0.9^{+0.3}_{-0.3}$	$0.21^{+0.08}_{-0.06}$
2.5	997	10.7±0.1	0.6±0.1	-21.8 ± 0.5	$0.41^{+0.09}_{-0.09}$	$-0.9^{+0.2}_{-0.2}$	...	...	$0.30^{+0.07}_{-0.08}$	$-0.9^{+0.2}_{-0.2}$	$0.38^{+0.11}_{-0.08}$
3.0	893	10.7±0.1	0.6±0.1	-21.9 ± 0.4	$0.4^{+0.1}_{-0.1}$	$-0.8^{+0.2}_{-0.3}$	...	...	$0.37^{+0.1}_{-0.09}$	$-0.7^{+0.2}_{-0.2}$	$0.4^{+0.1}_{-0.1}$
3.5	642	10.8±0.1	0.6±0.1	-22.0 ± 0.4	$0.5^{+0.2}_{-0.2}$	$-0.6^{+0.3}_{-0.4}$	...	...	$0.4^{+0.1}_{-0.1}$	$-0.7^{+0.3}_{-0.4}$	$0.4^{+0.2}_{-0.2}$
4.0	372	10.8±0.1	0.5±0.1	-22.0 ± 0.4	$0.8^{+0.2}_{-0.2}$	$-0.9^{+0.3}_{-0.3}$	...	...	$0.5^{+0.2}_{-0.2}$	$-1.0^{+0.3}_{-0.4}$	$0.7^{+0.2}_{-0.2}$
$11.0 < \log M_*/M_\odot] < 11.5$, $0.4 < z < 0.8$, $0.02 < R/R_{200} < 0.5$, $2.5 < R/R_h < 45$											
1.5	331	11.2±0.1	0.6±0.1	-22.8 ± 0.4	$0.5^{+0.1}_{-0.1}$	$-0.7^{+0.3}_{-0.2}$	> 0.7	< 54	$0.4^{+0.1}_{-0.1}$	$-0.8^{+0.3}_{-0.3}$	$0.4^{+0.2}_{-0.2}$
2.0	331	11.2±0.1	0.6±0.1	-22.8 ± 0.4	$0.6^{+0.2}_{-0.2}$	$-0.8^{+0.2}_{-0.3}$	...	...	$0.5^{+0.1}_{-0.1}$	$-0.8^{+0.3}_{-0.2}$	$0.5^{+0.2}_{-0.2}$
2.5	331	11.2±0.1	0.6±0.1	-22.8 ± 0.4	$1.0^{+0.2}_{-0.2}$	$-0.8^{+0.2}_{-0.2}$	...	...	$0.8^{+0.2}_{-0.2}$	$-0.9^{+0.2}_{-0.2}$	$0.9^{+0.2}_{-0.2}$
3.0	331	11.2±0.1	0.6±0.1	-22.8 ± 0.4	$1.2^{+0.2}_{-0.2}$	$-0.9^{+0.2}_{-0.2}$	> 0.8	< 16	$1.0^{+0.2}_{-0.2}$	$-0.9^{+0.2}_{-0.2}$	$1.1^{+0.3}_{-0.3}$
3.5	322	11.2±0.1	0.6±0.1	-22.8 ± 0.4	$1.7^{+0.3}_{-0.3}$	$-0.8^{+0.2}_{-0.2}$	> 0.7	< 21	$1.2^{+0.2}_{-0.2}$	$-0.8^{+0.2}_{-0.2}$	$1.6^{+0.3}_{-0.3}$
4.0	272	11.2±0.1	0.6±0.1	-22.8 ± 0.4	$1.8^{+0.3}_{-0.3}$	$-0.8^{+0.2}_{-0.2}$	> 0.8	< 28	$1.5^{+0.3}_{-0.3}$	$-0.7^{+0.2}_{-0.2}$	$1.7^{+0.3}_{-0.3}$
4.5	170	11.2±0.1	0.6±0.1	-22.8 ± 0.4	$2.4^{+0.5}_{-0.5}$	$-0.9^{+0.2}_{-0.2}$	> 1.0	< 18	$1.9^{+0.4}_{-0.5}$	$-0.9^{+0.2}_{-0.2}$	$2.3^{+0.5}_{-0.5}$
5.0	86	11.2±0.1	0.5±0.1	-22.8 ± 0.4	$2.3^{+0.7}_{-0.7}$	$-0.9^{+0.3}_{-0.3}$	> 0.8	< 26	$2.7^{+0.8}_{-0.8}$	$-0.7^{+0.3}_{-0.3}$	$2.1^{+0.7}_{-0.7}$
5.5	31	11.2±0.1	0.5±0.1	-22.9 ± 0.3	5^{+2}_{-1}	$-1.1^{+0.3}_{-0.3}$	> 0.7	< 23	4^{+2}_{-1}	$-0.9^{+0.3}_{-0.3}$	5^{+1}_{-1}

Table 3.2: Summary of early-type host galaxy properties and results from the inference using distances scaled by R_{200}

Δm	N_h	$\log[M_*/M_\odot]$ (a)	z(b)	M_r (a)	N_s (c)	γ_p (c)	A (c,d)	$\lvert\phi\rvert$ (c,d)	N_s (e)	γ_p (e)	Ns (c,f)
$10.5 < \log[M_*/M_\odot] < 11.0$, $0.1 < z < 0.4$, $0.07 < R/R_{200} < 1.0$, $4 < R/R_h < 45$											
1.5	274	10.7±0.5	0.31±0.07	-21.6 ± 0.5	...(g)	...	...	...	...	...	$0.12^{+0.1}_{-0.08}$
2.0	274	10.7±0.5	0.31±0.07	-21.6 ± 0.5	...	...	...	...	...	...	$0.2^{+0.2}_{-0.1}$
2.5	274	10.7±0.5	0.31±0.07	-21.6 ± 0.5	...	...	...	...	...	...	$0.3^{+0.2}_{-0.2}$
3.0	274	10.7±0.5	0.31±0.07	-21.6 ± 0.5	$0.2^{+0.3}_{-0.2}$	$-1.2^{+0.6}_{-0.9}$	< 1.3	...	$0.2^{+0.2}_{-0.1}$	$-1.7^{+0.9}_{-1}$	$0.3^{+0.2}_{-0.2}$
3.5	274	10.7±0.5	0.31±0.07	-21.6 ± 0.5	$0.3^{+0.3}_{-0.2}$	$-1.2^{+0.6}_{-0.8}$	< 1.3	...	$0.3^{+0.3}_{-0.2}$	$-1.4^{+0.7}_{-1}$	$0.4^{+0.3}_{-0.3}$
4.0	274	10.7±0.5	0.31±0.07	-21.6 ± 0.5	$0.5^{+0.3}_{-0.3}$	$-1.4^{+0.4}_{-0.6}$	< 1.3	...	$0.6^{+0.3}_{-0.3}$	$-1.4^{+0.4}_{-0.5}$	$0.7^{+0.3}_{-0.3}$
4.5	262	10.7±0.5	0.31±0.07	-21.6 ± 0.5	$0.8^{+0.4}_{-0.3}$	$-1.4^{+0.4}_{-0.5}$	...	...	$0.8^{+0.3}_{-0.4}$	$-1.4^{+0.4}_{-0.5}$	$0.9^{+0.4}_{-0.3}$
5.0	228	10.7±0.5	0.30±0.07	-21.7 ± 0.5	$0.8^{+0.5}_{-0.4}$	$-1.4^{+0.5}_{-0.6}$	< 1.3	...	$1.0^{+0.5}_{-0.5}$	$-1.1^{+0.4}_{-0.5}$	$1.1^{+0.5}_{-0.4}$
5.5	158	10.7±0.5	0.28±0.08	-21.7 ± 0.5	$1.1^{+1}_{-0.7}$	$-1.2^{+0.6}_{-1.0}$	< 1.3	...	$1.4^{+0.8}_{-0.8}$	$-1.1^{+0.4}_{-0.8}$	$1.4^{+0.7}_{-0.6}$
6.0	94	10.8±0.1	0.25±0.08	-21.8 ± 0.5	2^{+2}_{-1}	$-1.1^{+0.5}_{-0.9}$	< 1.3	...	2^{+1}_{-1}	$-1.0^{+0.6}_{-1}$	2^{+1}_{-1}
6.5	39	10.7±0.1	0.18±0.05	-21.6 ± 0.5	...	...	...	...	...	...	$1.0^{+1}_{-0.8}$
7.0	25	10.7±0.1	0.16±0.04	-21.6 ± 0.5	...	...	...	...	...	...	4^{+3}_{-2}
7.5	15	10.7±0.1	0.13±0.03	-21.4 ± 0.4	...	...	...	...	...	...	5^{+5}_{-3}
$11.0 < \log[M_*/M_\odot] < 11.5$, $0.1 < z < 0.4$, $0.07 < R/R_{200} < 1.0$, $4 < R/R_h < 45$											
1.5	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	...	...	...	...	...	...	$0.4^{+0.3}_{-0.2}$
2.0	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	...	...	...	...	...	...	$0.4^{+0.3}_{-0.3}$
2.5	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	...	...	...	...	...	...	$0.4^{+0.4}_{-0.3}$
3.0	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	$0.5^{+0.5}_{-0.3}$	$-1.7^{+0.8}_{-1}$	...	...	$0.8^{+0.5}_{-0.5}$	$-1.1^{+0.6}_{-0.8}$	$0.7^{+0.5}_{-0.4}$
3.5	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	$0.6^{+0.5}_{-0.4}$	$-1.8^{+0.7}_{-1}$	...	...	$0.8^{+0.6}_{-0.5}$	$-1.3^{+0.6}_{-0.9}$	$0.9^{+0.6}_{-0.5}$
4.0	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	$0.8^{+0.6}_{-0.5}$	$-1.8^{+0.6}_{-0.9}$	...	...	$1.1^{+0.7}_{-0.6}$	$-1.3^{+0.6}_{-0.8}$	$1.2^{+0.6}_{-0.6}$
4.5	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	$0.8^{+0.6}_{-0.5}$	$-2.3^{+0.7}_{-1}$	...	...	$1.1^{+0.8}_{-0.6}$	$-1.8^{+0.6}_{-0.9}$	$1.4^{+0.7}_{-0.7}$
5.0	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	$1.0^{+0.6}_{-0.5}$	$-2.5^{+0.7}_{-1}$	...	...	$1.0^{+0.6}_{-0.5}$	$-2.3^{+0.7}_{-0.9}$	$1.7^{+0.7}_{-0.8}$
5.5	54	11.2±0.1	0.32 ±0.06	-22.3 ± 0.4	$1.2^{+0.8}_{-0.5}$	$-2.3^{+0.7}_{-0.8}$	...	...	$1.4^{+1}_{-0.7}$	$-1.8^{+0.6}_{-0.8}$	$2.0^{+0.9}_{-1}$
6.0	45	11.2±0.1	0.31±0.06	-22.4 ± 0.4	2^{+2}_{-1}	$-1.9^{+0.6}_{-0.7}$	...	...	2^{+2}_{-1}	$-1.5^{+0.5}_{-0.8}$	3^{+1}_{-1}
6.5	14	11.2±0.1	0.24±0.04	-22.3 ± 0.5	...	...	...	...	...	...	2^{+3}_{-2}
7.0	9	11.2±0.1	0.23±0.04	-22.4 ± 0.5	...	...	...	...	...	...	4^{+4}_{-3}
7.5	9	11.2±0.1	0.23±0.04	-22.4 ± 0.5	...	...	...	...	...	...	7^{+6}_{-4}
$10.5 < \log[M_*/M_\odot] < 11.0$, $0.4 < z < 0.8$, $0.07 < R/R_{200} < 1.0$, $4 < R/R_h < 45$											
1.5	857	10.7 ±0.1	0.6±0.1	-21.8 ± 0.5	$0.2^{+0.2}_{-0.1}$	$-1.0^{+0.4}_{-0.5}$	< 1.0	...	$0.13^{+0.08}_{-0.08}$	$-1.2^{+0.5}_{-0.6}$	$0.23^{+0.1}_{-0.09}$
2.0	857	10.7 ±0.1	0.6±0.1	-21.8 ± 0.5	$0.2^{+0.1}_{-0.1}$	$-1.2^{+0.5}_{-0.8}$	< 1.1	...	$0.09^{+0.08}_{-0.07}$	$-1.5^{+0.7}_{-1}$	$0.2^{+0.1}_{-0.1}$
2.5	814	10.7 ±0.1	0.6±0.1	-21.8 ± 0.5	$0.2^{+0.2}_{-0.1}$	$-1.3^{+0.5}_{-0.7}$	< 1.1	...	$0.16^{+0.1}_{-0.07}$	$-1.6^{+0.5}_{-0.6}$	$0.3^{+0.1}_{-0.1}$
3.0	726	10.7 ±0.1	0.6±0.1	-21.9 ± 0.5	$0.4^{+0.2}_{-0.2}$	$-1.4^{+0.4}_{-0.5}$	< 1.1	...	$0.2^{+0.1}_{-0.1}$	$-1.8^{+0.5}_{-0.7}$	$0.5^{+0.2}_{-0.1}$
3.5	583	10.7 ±0.1	0.6±0.1	-22.0 ± 0.5	$0.5^{+0.2}_{-0.2}$	$-1.6^{+0.3}_{-0.5}$	< 0.7	...	$0.3^{+0.2}_{-0.2}$	$-1.7^{+0.4}_{-0.6}$	$0.6^{+0.2}_{-0.2}$
4.0	372	10.7 ±0.1	0.5 ±0.1	-22.1 ± 0.5	$0.5^{+0.3}_{-0.2}$	$-1.7^{+0.5}_{-0.6}$	< 1.1	...	$0.4^{+0.2}_{-0.2}$	$-1.9^{+0.5}_{-0.7}$	$0.7^{+0.2}_{-0.2}$
$11.0 < \log[M_*/M_\odot] < 11.5$, $0.4 < z < 0.8$, $0.07 < R/R_{200} < 1.0$, $4 < R/R_h < 45$											
1.5	182	11.2±0.1	0.6±0.1	-22.5 ± 0.5	$0.1^{+0.1}_{-0.1}$	$-2.0^{+0.8}_{-1}$	< 1.2	...	$0.2^{+0.1}_{-0.1}$	$-1.9^{+0.7}_{-1}$	$0.2^{+0.1}_{-0.1}$
2.0	182	11.2±0.1	0.6±0.1	-22.5 ± 0.5	$0.3^{+0.2}_{-0.2}$	$-1.8^{+0.5}_{-0.7}$	< 1.2	...	$0.4^{+0.2}_{-0.2}$	$-1.6^{+0.4}_{-0.6}$	$0.5^{+0.2}_{-0.2}$
2.5	182	11.2±0.1	0.6±0.1	-22.5 ± 0.5	$0.5^{+0.2}_{-0.2}$	$-2.1^{+0.5}_{-0.8}$	< 0.8	...	$0.6^{+0.2}_{-0.2}$	$-1.7^{+0.4}_{-0.5}$	$0.8^{+0.3}_{-0.2}$
3.0	181	11.2±0.1	0.6±0.1	-22.5 ± 0.5	$0.9^{+0.3}_{-0.3}$	$-1.6^{+0.4}_{-0.5}$	< 0.9	...	$1.0^{+0.3}_{-0.3}$	$-1.4^{+0.3}_{-0.3}$	$1.0^{+0.3}_{-0.3}$
3.5	170	11.2±0.1	0.6±0.1	-22.6 ± 0.4	$0.7^{+0.4}_{-0.3}$	$-1.7^{+0.5}_{-0.6}$	< 1.0	...	$0.8^{+0.4}_{-0.4}$	$-1.5^{+0.4}_{-0.7}$	$1.0^{+0.4}_{-0.3}$
4.0	139	11.2±0.1	0.6±0.1	-22.6 ± 0.4	$0.7^{+0.5}_{-0.4}$	$-1.6^{+0.5}_{-0.7}$	< 1.1	...	$1.0^{+0.5}_{-0.5}$	$-1.0^{+0.5}_{-0.7}$	$0.9^{+0.4}_{-0.4}$

Table 3.3: Summary of late-type host galaxy properties and inference results. (a) Geometric Means, (b) Mean, (c) Satellite distances scaled by R_{200}, (d) 68 % one-sided confidence interval, (e) Satellite distances scaled by R_h (f) Satellite distances scaled by R_h, Inference performed with a Gaussian prior on γ_p with mean -1.1 and standard deviation 0.3, (g) Blank spaces indicate no inference on the parameter as the posterior PDF was essentially uniform

3.12.1 Source Extractor Parameters and Photometry Comparison

After removing host galaxy light from cutout images, we run SExtractor on the residual image to identify remaining objects. In this step we use SExtractor parameters tuned to match the photometry and completeness of the COSMOS photometric catalog. We could not directly the parameters listed by the COSMOS team, as their photometry was performed on images with $0.''03$ resolution pixels while at the time of this work, the publicly available images had $0.''05$ pixels.

To check the effects of using different SExtractor parameters and pixel scales, we compare our completeness and photometry to that of the COSMOS photometric catalog in a 4 arcmin^2 cutout of the COSMOS field. In this area, we detect 817 objects with I-band MAG_AUTO<25. Of the 817 objects we identified, 792 have matches in the COSMOS photometric catalog to the same depth and within half an arcsecond. In Figure 3.10, we compare the MAG_AUTO output from SExtractor to the values in the COSMOS photometric catalog for matching objects. The mean difference in the MAG_AUTO estimate is $(-2.53 \pm 51) \times 10^{-3}$. The major outliers in the MAG_AUTO comparison are all in areas of high object density and thus most likely due to differences in deblending.

A full list of our SExtractor parameters is given in Table 3.4

Parameter	Value		
	Large Object Mask	Point Object Mask	Final Photometry
DETECT_MINAREA	10	6	6.5
DETECT_THRESH	1.8	3	1.7
ANALYSIS_THRESH	1.8	3	1.7
DEBLEND_NTHRESH	64	64	64
DEBLEND_MINCONT	1E-6	1E-6	0.05
FILTER_NAME	gauss_2.5_5x5.conv	gauss_1.5_3x3.conv	gauss_2.5_5x5.conv
BACK_TYPE	MANUAL		
BACK_VALUE	0.0		
SEEING_FWHM	0.1		
INTERP_MAXXLAG	2		
INTERP_MAXYLAG	2		
WEIGHT_TYPE	MAP_RMS		

Table 3.4: SExtractor Parameters. Parameters not listed are SExtractor default values. Blank spaces indicate that the parameter is the same for all rows.

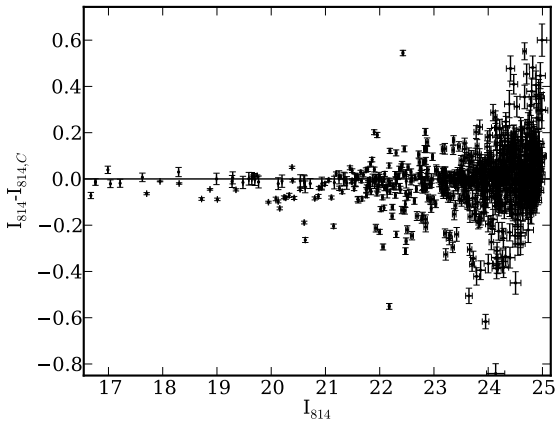

Figure 3.10: Comparison of our photometry with no host galaxy subtraction to COSMOS photometry in an 18 arcmin ^{2} field cutout.

3.13 Completeness

To estimate our sensitivity to low surface brightness companions to host galaxies, we simulate dwarf galaxy-like objects in the ACS images with Sersic profiles given by de Rijcke et al. (2009) at a range of redshifts representative of the redshift range studied in this work. We then perform the host subtraction routine and estimate the fraction

of objects we recover and the output magnitude relative to the input magnitude. The results of these simulations are shown in Figure 3.11. For comparison we also show the results without host subtraction but with object detection performed using the same final SExtractor parameters as are used in the host-subtracted case.

We choose the minimum radius at which we search for satellites where completeness begins to drop below 90% for early-type hosts. For late-type hosts, the minimum radius is chosen to ensure spiral arms are excluded from the analysis. Images after host subtraction are checked for every host to ensure that these minimum radii exclude spiral arms or extended residuals from disk structure (see Chapter 2).

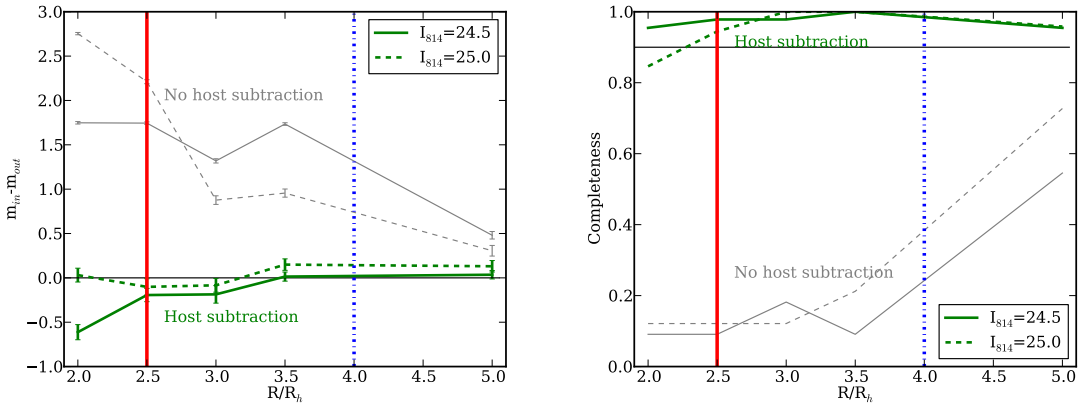

Figure 3.11: Results of completeness simulations. Red solid and blue dot-dashed vertical lines show the imposed minimum radii at which we study the satellite population for early and late-type galaxies respectively. On the left, the horizontal black line indicates that the input and output magnitudes are the same. On the right the horizontal line indicates 90% completeness. Host subtraction is essential both for photometric accuracy and for completeness at small radii.

3.13.1 Full Posterior PDFs

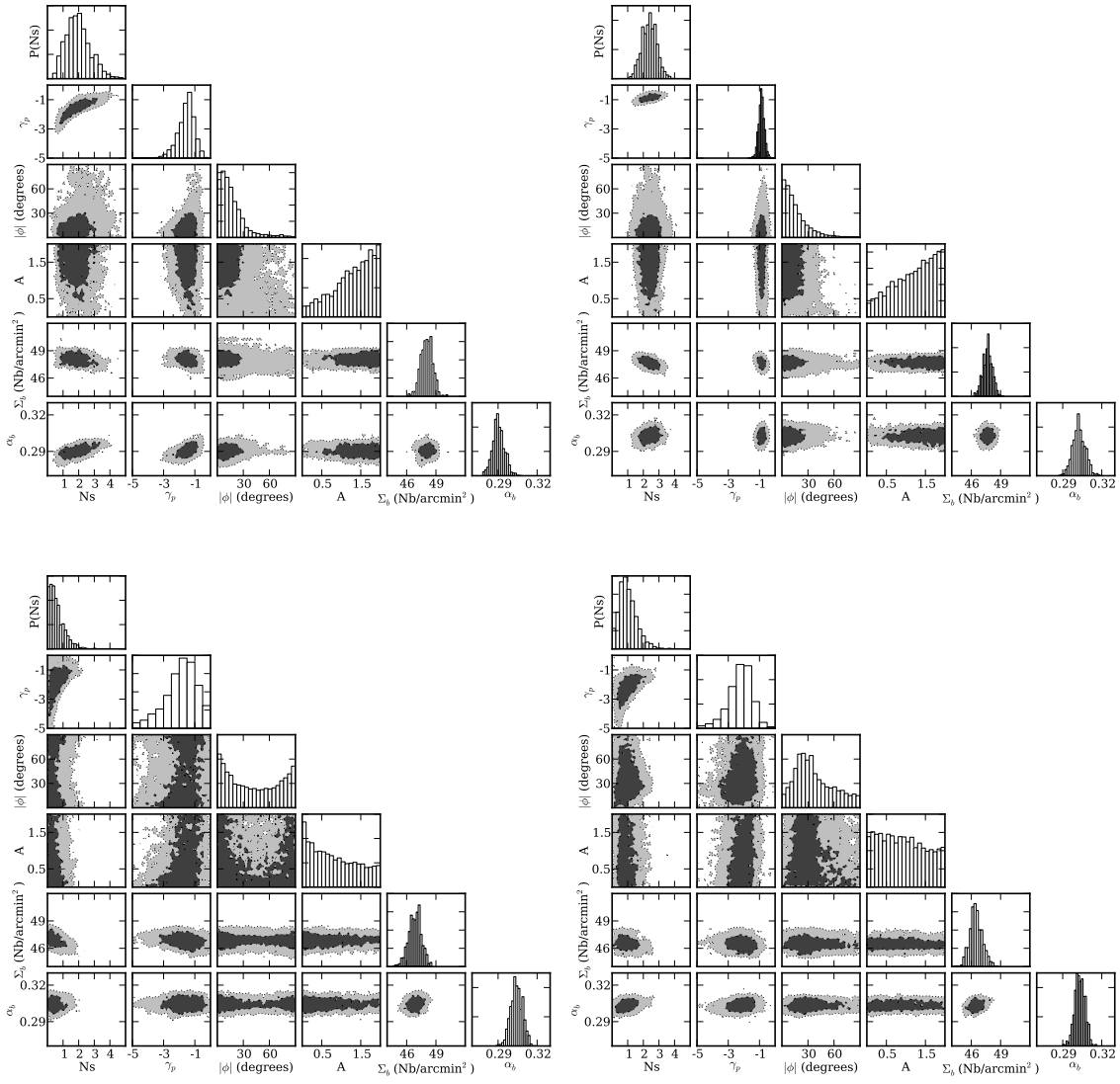

Figure 3.12: Full posterior PDFs for dm = 4.0 satellite of high mass *clockwise from top left*: Low redshift early-type hosts, high redshift early-type hosts, high redshift late-type hosts, low redshift late-type hosts. Dark and light gray contours indicate regions of 68 and 95% confidence respectively.

Chapter 4

Luminous satellites III: Comparison of the luminosity function with theoretical predictions

This chapter was published as Nierenberg, A. M., Treu, T., Menci, N., Lu, Y., Wang, W. "The Cosmic Evolution of Faint Satellite Galaxies as a Test of Galaxy Formation and the Nature of Dark Matter", ApJ, 772, 146, (2013) and is included here with minor adjustments

As was discussed in the Introduction, star formation in satellite galaxies is not yet well understood because it is driven by a large number of non-linear processes. A number of astrophysical models, with varying models of star formation can all reproduce the luminosity function of satellites around low redshift, Milky Way mass host galaxies. In this Chapter, we demonstrate how the observations from the previous two chapters make it possible to break degeneracies in these star formation models, by testing the prediction in a new range of redshifts and host stellar masses.

We focus on three semi-analytic galaxy formation models which we discuss in more detail below. We also implement one of the models with both Warm and Cold Dark Matter substructure mass function in order to show how variations in the halo mass function produce different predictions from variations in baryonic physics. We compare

the predictions from these models with the observed abundance of satellite galaxies as a function of host galaxy mass and cosmic time.

This test is made possible by two recent developments. On the observational side, the implementation of powerful statistical tools to detect and count satellites in deep archival Hubble Space Telescope images provides data to compare with model predictions as discussed in Chapters 2, and 3. On the theoretical side, advances in computational methods have allowed for cosmological simulations with unprecedented volume and resolution. In this work we focus on predictions from four independent cosmological simulations, which have semi-analytic models applied to dark matter merging trees, which we describe in more detail below.

This chapter is organized as follows: In §4.1 we summarize the key aspects of the theoretical models. In §4.2, we compare the observations with the theoretical predictions. In §4.3 we compare a few of the main properties of the models and present new predictions for the distribution of colors of satellite galaxies across cosmic time. In §4.4 we compare the model results with other observations taken from the literature. Finally in §4.5 we conclude with a discussion and summary of the results.

4.1 Theoretical Models

We study three different semi-analytic models implemented in three different cosmological Cold Dark Matter simulations. In addition, we apply one of the semi-analytic models to a Warm Dark Matter cosmology. This section provides a description of both the Cold and Warm Dark Matter simulations and their models for galaxy formation.

4.1.1 Cold Dark Matter Models

When studying the satellite luminosity function, it is extremely useful to compare multiple cosmological simulations along with different star formation parameters in order to understand the range of theoretical predictions given different assumptions. In this work we focus on three Cold Dark Matter cosmological simulations, each with its own semi-analytic model for galaxy. The simulation parameters governing star formation and the underlying dark matter mass function were fully specified prior to performing this test and are the same as in previous papers, in which a more complete description can be found (Menci et al. 2012; Lu et al. 2012; Guo et al. 2011b). Below we briefly summarize the relevant aspects of these simulations.

Menci et al. 2012

The semi-analytic models are described in detail by Menci et al. (2012). In brief, the backbone of the models are computations of dark matter merging histories that can resolve halos down to masses of $10^5 M_\odot$, allowing for an extremely accurate characterization of the survival and merger histories of satellites. This is essential for studying the faint end of the satellite luminosity function. The models predict mass distributions of halos, and the merging rates of DM halos and subhalos based on an input power spectrum, and generate luminosities for these halos based on a set of star formation prescriptions.

These galaxy formation prescriptions connect the complex baryon physics of star formation to the dynamical evolution of the DM haloes. For each subhalo hosting a galaxy, the model predicts radiative gas cooling, ensuing star formation and associated

feedback from Supernovae (SNae) events. Cooled gas settles into a rotationally supported disk with radius and rotational velocity related to the DM mass of the subhalo. This gas gradually condenses into stars at a rate consistent with the observed Kennicut-Schmidt law (Kennicutt 1998). SNae return part of the cooled gas to the hot gas phase at the virial temperature of the halo. Star formation is also triggered by galaxy-galaxy interaction-driven starbursts, which induce gas accretion onto the central supermassive Black Holes.

It is important to note that the set of parameters used in these models is much smaller than the number of observables they are simultaneously consistent with. In brief, the star formation parameters are tuned to simultaneously match the field luminosity function in multiple bands from redshifts 0 to 6, the Tully-Fisher relation (Tully & Fisher 1977), the stellar mass function of field galaxies from between redshifts 0 and 4, the stellar mass-star formation relation to a redshift of 2, the colors and color magnitude relations of both field and cluster galaxies in addition to the AGN luminosity function.

The observations from the previous chapter required host galaxies to be outside of R_{200} of any galaxy with a higher stellar mass, where R_{200} was estimated using the formula from Dutton et al. (2010). Ideally this selection would be duplicated when comparing with the simulation, however this is not possible for the Menci et al. (2012) models, as they do not contain spatial information. This should not significantly affect our results, as comparisons with the Millennium and Bolshoi N-Body simulations (described below) show that the central isolation criteria used in the previous chapters are efficient at returning central host galaxies ($\sim 90\%$), and that the use of the matching selection criteria in these cases did not significantly alter the inferred predicted number of satellites per host, relative to simply using the simulation selection of central galaxies.

Guo et al. 2011

The semi-analytic model of Guo et al. (2011b) is based on two very large dark matter simulations, the Millennium Simulation (MS; Springel et al. 2005) and the Millennium-II Simulation (MS-II; Boylan-Kolchin et al. 2009). The box size of the MS is $500h^{-1}$Mpc and its merger trees are complete for subhaloes above a mass limit of $1.7 \times 10^{10}h^{-1}\mathrm{M}_{\odot}$. The MS-II follows a cube of side $100h^{-1}$Mpc, but with 125 times better mass resolution than the MS (subhalo masses greater than $1.4 \times 10^{8}h^{-1}\mathrm{M}_{\odot}$). Both simulations adopt the WMAP1-based ΛCDM cosmology (Spergel et al. 2003) with parameters $h = 0.73, \Omega_m = 0.25, \Omega_\Lambda = 0.75, n = 1$ and $\sigma_8 = 0.9$. Due to the lower resolution of the MS, Wang & White (2012) found that the luminosity functions of satellites flattens for satellites with $M_r > -18$. For brighter satellites with $M_r < -18$, the simulations are consistent. Since satellite luminosity functions will be measured to about three-orders of magnitude fainter than central primaries in this chapter, we will focus on the semi-analytic model implemented on MS-II hereafter[1].

In general, the galaxy evolution model of Guo et al. (2011b) is based on those developed by Springel et al. (2005); Croton et al. (2006); De Lucia & Blaizot (2007). The model includes a few main modifications such as the different definition of satellite galaxies, the gradual stripping and disruption of satellites, a mass-dependent model of supernova feedback, a modified model for reionization and a more realistic treatment of the growth of stellar and gaseous disks. Free parameters of these models by (Guo et al. 2011b) were determined to give close predictions to the abundance and clustering of low redshift

[1]Here we use the data downloaded from http://www.mpagarching.mpg.de/millennium for registered users.

galaxies, as inferred from SDSS, and are functions of their stellar mass, luminosity and color.

There are two types of satellites in the simulation: *i*) those with an associated dark matter subhalo (type-1) and *ii*) those whose dark subhalo has fallen below the resolution limit of the simulation (type-2). For the latter, the position and velocity of the orphan galaxy is given by those of their most bound particle. Type-2 satellites are removed from the galaxy catalogues when one of these two conditions is fulfilled: 1) the time passed from the disruption of the subhalo is longer than their estimated dynamical-friction timescale, or 2) the integrated tidal forces from the host halo exceed the binding energy of the galaxy.

We project the simulation box in three orthogonal directions (along their x, y and z axes). In each projection we assign each galaxy a redshift based on its "line-of-sight" distance and peculiar velocity. We select isolated primaries using criteria which are directly analogous to those used to obtain the observational results in the previous chapters. Satellites are defined to be all companion galaxies whose distances to the isolated primaries are smaller than the halo virial radius (R_{200}). To directly compare results based on COSMOS, the real R_{vir} provided in the database is ignored, and instead we used the empirical formula of Dutton et al. (2010) to estimate M_{vir} and R_{200} from stellar masses of galaxies.

Lu et al. 2012

The baryonic processes implemented in this semi-analytic model are described in Lu et al. (2011, 2012). Different from the previously published versions, this version of the

model is applied on a set of halo merger trees extracted from a large cosmological N-body simulation, the Bolshoi simulation (Klypin et al. 2011), which has a box size $250h^{-1}$ Mpc on a side. The simulation adopted a cosmology favored by WMAP7 data (Jarosik et al. 2011) and WMAP5 data (Dunkley et al. 2009; Komatsu et al. 2009) with parameters with $\Omega_{\rm m,o} = 0.27$, $\Omega_{\Lambda,o} = 0.73$, $\Omega_{\rm b,o} = 0.044$, $h = 0.70$, $n = 0.95$ and $\sigma_8 = 0.82$. The mass resolution of the simulation is $1.35 \times 10^8 h^{-1}{\rm M}_\odot$, which allows us to track halos and subhalos with mass $\sim 7 \times 10^9 h^{-1}{\rm M}_\odot$. Dark matter halos and subhalos are identified with the Rockstar halo finder (Behroozi et al. 2012) based on adaptive hierarchical refinement in phase-space.

As with the two other models in this work, the semi-analytic model follows the dark matter merger tree and calculates the rates of gas cooling, star formation, outflow induced by star formation feedback, and galaxy-galaxy mergers. The kinematics of satellite galaxies is followed by using subhalo information from the simulation whenever the subhalo is resolved. When a halo becomes a subhalo, we instantaneously strip the hot gas associated with the halo, while the stellar mass and cold gas mass remain intact.

When the subhalo is no longer resolved in the simulation, the model applies dynamical friction (Binney & Tremaine 1987) to estimate when the satellite has merged into the central galaxy, and it assumes that the tidal stripping is strong enough to also strip the stellar mass and cold gas mass. At this point, the entire cold gas disk is stripped and is mixed into the hot gas of the host primary halo. Starting from that time, a fraction of stellar mass is tidally stripped per orbital timescale. The efficiency of the tidal stripping is controlled by a parameter, which is tuned to yield 30% of the stellar mass is stripped in every orbital timescale to match the conditional stellar mass function of local galaxies.

Other parameters governing star formation and feedback are tuned using an MCMC optimization to match the local galaxy stellar mass function (Moustakas et al 2013). The model is guaranteed to produce a global galaxy stellar mass function which provides the best possible match to the data between 10^9 and 10^{12} $M_\odot$ at redshift zero, within the observational uncertainty and given the chosen model parametrization.

To compare with observation, hosts were selected using the same isolation criteria as used for the observations, with hosts required to be not within R_{200} of a host with higher stellar mass, where R_{200} is estimated from Dutton et al. (2010), using the relationship for early-type hosts. The satellite luminosity function was measured within this region, rather within R_{vir} as given by the simulation.

4.1.2 Menci Warm Dark Matter Model

Warm Dark Matter has been proposed as a means of suppressing the satellite luminosity function by reducing the number of low mass dark matter halos, suppressing power with respect to CDM below a certain cutoff-scale.

In this work, we use WDM merger trees from Menci et al. (2012) which are based on a cutoff scale of $\sim$ 1 Mpc, corresponding to a $\sim$0.75 keV/c^2 thermal relic, which suppresses the power spectrum on sub-Mpc scales and has behavior equivalent to that of a $\sim$ 3.4 keV/c^2 sterile neutrino. We chose this mass to be low enough to affect satellite-scale structure while still agreeing with limits from observations of large-scale and local group structure which constrain the particle mass to be larger than 0.6 keV/c^2 for a thermal relic and 2.5 keV/c^2 for a sterile neutrino DM particle (Viel et al. 2009; Boyarsky et al.

2009; Polisensky & Ricotti 2011; Kang et al. 2012).

The fiducial mass function from Menci et al. (2012) was based on an Extended Press Schechter formalism (EPS, Bond et al. 1991). In our predictions, we consider the effects of a complete suppression of progenitors with masses below that corresponding to the free-streaming scale; this maximizes the possible effects of different window functions and collapse thresholds in the WDM merging trees (see Benson et al. 2013). We include the possible range of model predictions under different assumptions for the building up of merging trees in the uncertainty regions in Figure 4.1.

The mass function from Menci et al. (2012) is based on on the Extended Press Schechter formalism (EPS, Bond et al. 1991), which is modified to take into account the suppression of structure below the free-streaming scale (see, e.g. Menci et al. 2012; Benson et al. 2013). The uncertainties associated with this modification are reflected in the width of the prediction for the final WDM satellite luminosity function.

Our goal in this comparison was to explore how the effects of varying the power spectrum compare with the effects of varying semi-analytic star formation parameters. To achieve this, *the same* semi-analytic model of star formation is used for the WDM models as was applied to the CDM merging trees in the Menci CDM model. The CDM semi-analytic model was selected to provide a good match to the field color distribution. Menci et al. (2012) show that this model provides a good fit to the field luminosity function when applied to a WDM power spectrum.

4.2 Results

The improvements in observations and the implementation of semi-analytic models in cosmological scale simulations allows a comparison between the observed and predicted number of satellites as a function of host galaxy mass and cosmic time for the first time, thus allowing for significantly more discriminatory power than tests based only on the MW or the local volume.

The results of the comparison are presented in Figure 4.1, where we plot the observed number of satellites as a function of the ratio between host and satellite luminosities in SDSS-r, along with the WDM and CDM model predictions. Among the CDM model predictions, two main trends are evident. First, no one CDM prediction precisely matches the observation at all redshift and stellar mass intervals. This highlights the importance of using data from a range of redshift and stellar mass when tuning the parameters of semi-analytic models. All three models show qualitatively similar behavior with a strong dependence on the number of satellites per host on the host galaxy mass. Taking into account the covariance between data points (calculated by bootstrap resampling), the generalized chi-squared between the models and the data is 529, 653 and 105 for the Menci, Guo and Lu models respectively for 41 degrees of freedom.

The Menci WDM model shows distinct behavior in comparison with the CDM models. Notably, it predicts weaker host mass dependence and less redshift evolution than predicted by any of the CDM models. Of the four models it provides the best agreement with the data with a chi-squared of 56.

One of the most interesting results of the above comparison is the significant difference

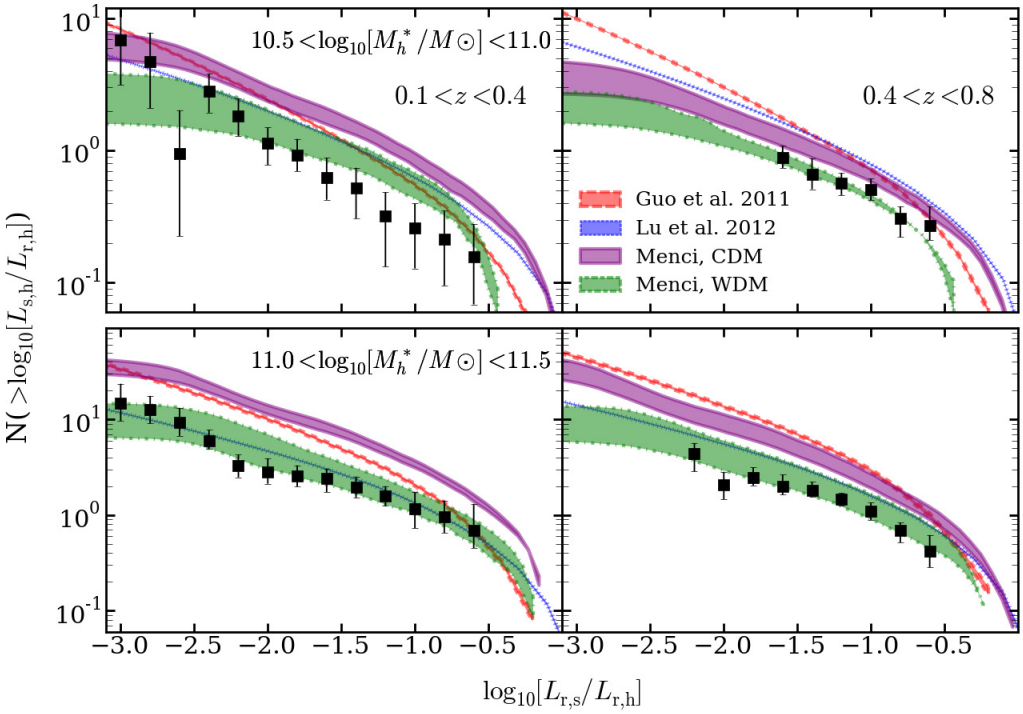

Figure 4.1: Comparison of CDM and WDM model predictions to observed number of satellites over a range of redshift and host stellar masses. The purple, blue and red curves represent predictions from three CDM based simulations with separate semi-analytic models for star formation. The gray curve shows the prediction for one WDM simulation with star formation parameters identical to those used in Menci, CDM prediction. For the Lu and Guo models, the line widths represent the scatter in the mean value across the simulation boxes. For the Menci CDM model, the shaded region represents the range of model predictions possible given the observational uncertainty in the selection of host halo masses. For the Menci WDM model, the shaded region also accounts for the uncertainties in the suppression of progenitors below the free-streaming scale (see 4.1.2) The points with vertical error bars are measurements from the previous chapter. The top panels show the comparison for lower mass hosts at lower (left) and higher redshift (right), while the lower panels are the equivalent for higher mass hosts.

in the predicted satellite luminosity functions, even among the three CDM models. In the following section, we perform a detailed comparison of some of the properties of the four models in order to explore the cause of these differences.

4.3 Comparison of simulation properties

In the previous section we found significant differences in the model predictions for the satellite luminosity function. As discussed in Section 4.1, numerous physical processes contribute to the final predicted luminosity function. In this section, we compare key aspects of the models in order to elucidate which model assumptions drive the predicted differences. We first compare the subhalo mass function in the models, before the effects of tidal stripping take effect, for all subhalos regardless of final luminosity. We then show the stellar mass to halo mass relationship for hosts and satellites in the luminosity range considered in this work, before the subhalos have undergone tidal stripping. We conclude by comparing the predicted satellite colors to show the differences in star formation histories.

4.3.1 Unstripped subhalo mass function

The four simulations all assumed slightly different cosmologies and had different mass resolutions, so it is important to consider how much the differences in the predicted results are due to these effects. In Figure 4.2 we plot the mass function of subhalos in a mass range representative of the hosts and subhaloes from Figure 4.1. Note that this mass function is purely based on dark matter mass, and does not directly correspond to the

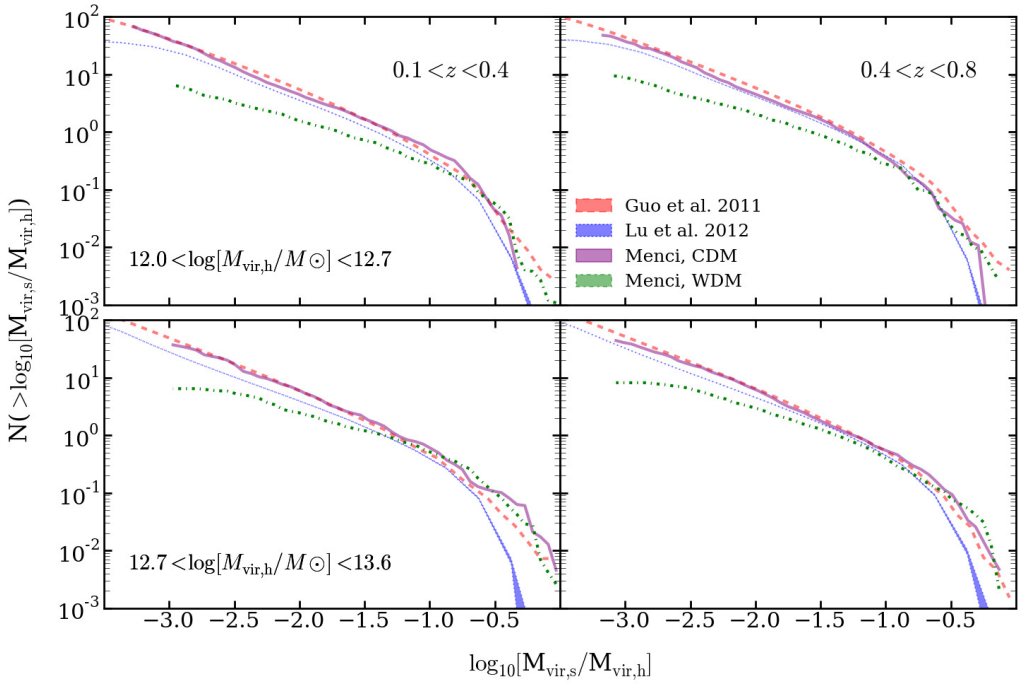

Figure 4.2: Comparison of the subhalo mass function at the time of subhalo accretion for the four models, for a range of masses representative of the host halos and subhalos in Figure 4.1. Note that these mass functions do not include the effects of tidal stripping by the main halo, nor are they selected to host luminous satellite galaxies, or host galaxies with the same stellar masses as the hosts in Figure 4.1

satellite galaxies in Figure 4.1, as there is no luminosity selection.

The Guo, Menci, and Lu CDM subhalo mass functions are extremely similar, while the WDM model shows significant truncation for subhalos with $\log[M_{sub}/M_{host}] < -1$ for the less massive hosts and $\log[M_{sub}/M_{host}] < -1.5$ for the more massive hosts, and predicts similar behavior to other WDM models with sub-Mpc cutoff scales (e.g. Colín et al. 2000; Smith & Markovic 2011; Kamada et al. 2013).

4.3.2 The halo mass to stellar mass relationship

One of the most important physical processes relied on in this comparison between simulation and observation, is the relationship between host halo mass and stellar mass, because the number of subhalos around a given host galaxy is strongly dependent on the halo mass of the host galaxy (see e.g. Busha et al. 2011), particularly in the case of CDM. Furthermore, as we always consider the quantity $\log_{10}[L_s/L_h] > -3$, it is important to distinguish whether differences in the models are caused by differences in L_s or in L_h.

In Figure 4.3 we plot the halo to stellar mass relationship for the four models, in addition to the observed relationship from Dutton et al. (2010), which was used to estimate R_{200} for the host galaxies in the observations. For bright galaxies with stellar masses greater than $10^{10.5}M_\odot$, the models are all very similar to each other and to the observed relationship, within the large scatter. Thus we conclude that the differences in amplitude in the predictions for the satellite luminosity function are not driven primarily by differences in the halo to stellar mass relationship for host galaxies.

Towards the faint end, all models are very similar within the scatter, although the Menci models show a marginally higher dark matter mass at fixed stellar mass. As we show in the following section, the Menci satellite colors tend to be bluer than the Guo and Lu colors. From (Bell et al. 2003), Table 7, the bluer colors of the Menci model, correspond to an r band mass-to-light ratio which is lower by roughly 0.1-0.2 dex on average than for the Guo and Lu models. Thus, as the final luminosity functions are similar, it is expected that the Menci models would produce galaxies with on average slightly lower stellar masses for fixed halo mass.

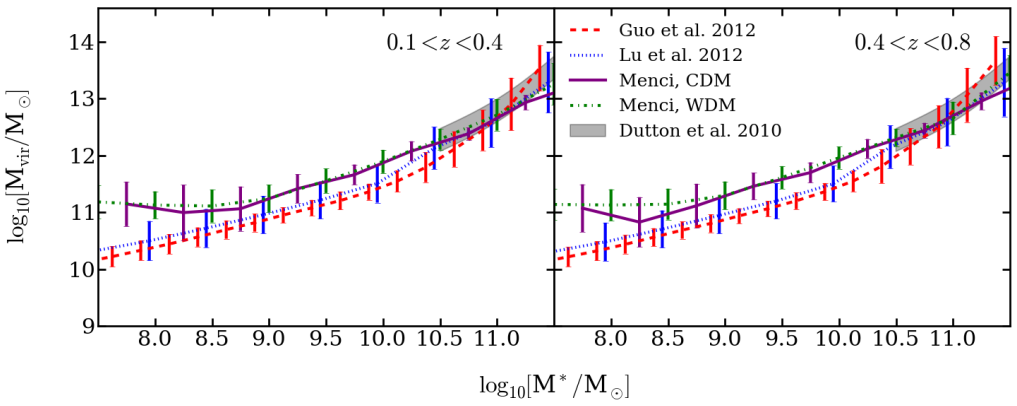

Figure 4.3: Comparison of the virial to stellar mass relationship for both host and satellite galaxies for the four models. The subhalo virial masses are taken at the time of accretion, and thus do not include the effects of tidal stripping.

4.3.3 Satellite galaxy colors

The intrinsic colors of satellite galaxies are dependent on the star formation history. As discussed above, this is determined by the metallicity, feedback and UV heating, in addition to environmental effects as the satellites enter the influence of the host galaxy halo. All three models used the same stellar population synthesis models from (Bruzual & Charlot 2003), thus satellite galaxy colors provide an important means of distinguishing between different physical models for the suppression of star formation in low mass halos. Two models which produce similar luminosity functions may rely on very different star formation prescriptions, which will result in different color distributions. In Figure 4.4, we plot the predicted distribution of rest-frame u-i colors for satellites with $\log_{10}[L_s/L_h] > -3$.

In order to summarize these predictions, we fit each of the distributions to a skewed normal distribution (Azzalini 1985) where the probability of having color c is given by:

$$P(c) \propto e^{\frac{(c-c_o)^2}{2\sigma^2}} \left(1 + erf\left(\frac{a(c-c_o)}{\sqrt{2}\sigma}\right)\right) \tag{4.1}$$

Here c_o and σ are the usual mean and standard deviations of a normal distribution while the parameter a describes the skewness. The best fit values for these parameters are listed in Table 4.1 [2]. The parametrization of the distributions as being skewed normal is not physically motivated, but rather intended to facilitate future comparison between observation and these predictions.

The models all predict significantly different color distributions for the satellite galaxies, with the exception of the Menci CDM and WDM models, for which the same star

[2]We do not consider the secondary peaks in the Menci models in this fit.

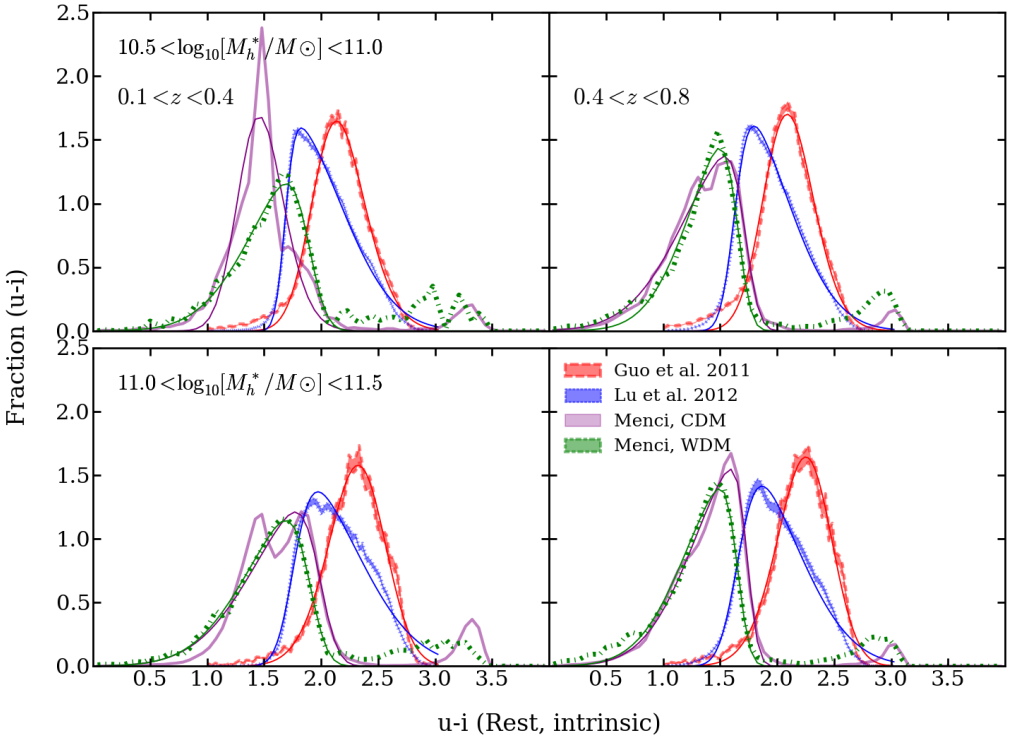

Figure 4.4: The normalized rest-frame distribution of colors, for the satellite galaxies plotted in 4.1. The solid lines indicate the best fit skewed normal distributions defined by Equation4.1, with fit parameters listed in Table 4.1

formation parameters were used. We highlight the fact that although the luminosity function predicted by the Menci WDM and Lu CDM models are similar within some redshift and host stellar mass ranges, the color distributions are very different. This is due to the fact that in the Lu CDM model, the faint end of the luminosity function is suppressed mostly by the effects of feedback and ram pressure stripping and heating of the gas by the host halo, while in the Menci WDM model, the luminosity function is suppressed by the lack of low mass subhalos, as expected.

In order to facilitate comparison with observation, it is important to also consider the effects of dust extinction. In Figure 4.5, we show the prediction for the rest-frame colors of satellite galaxies with dust. The Menci models behave very differently with the addition of dust, relative to the Guo and Lu models, with the Menci color distributions becoming much wider and significantly redder while the Guo and Lu models develop a longer redward tail, without significant other alteration. The secondary peaks in the Menci model become more prominent with the addition of dust, so we fit the Menci models with the sum of two skewed normal distributions with parameters reported in Table 4.1, where the value A_2/A_1 describes the relative amplitudes between the two skewed normals.

4.4 Comparison with literature

The models presented in this work predict a broad range of observables, many of which have already been compared to observations, either at lower redshifts, or of brighter objects than those considered in this chapter. Below we provide a brief comparison with

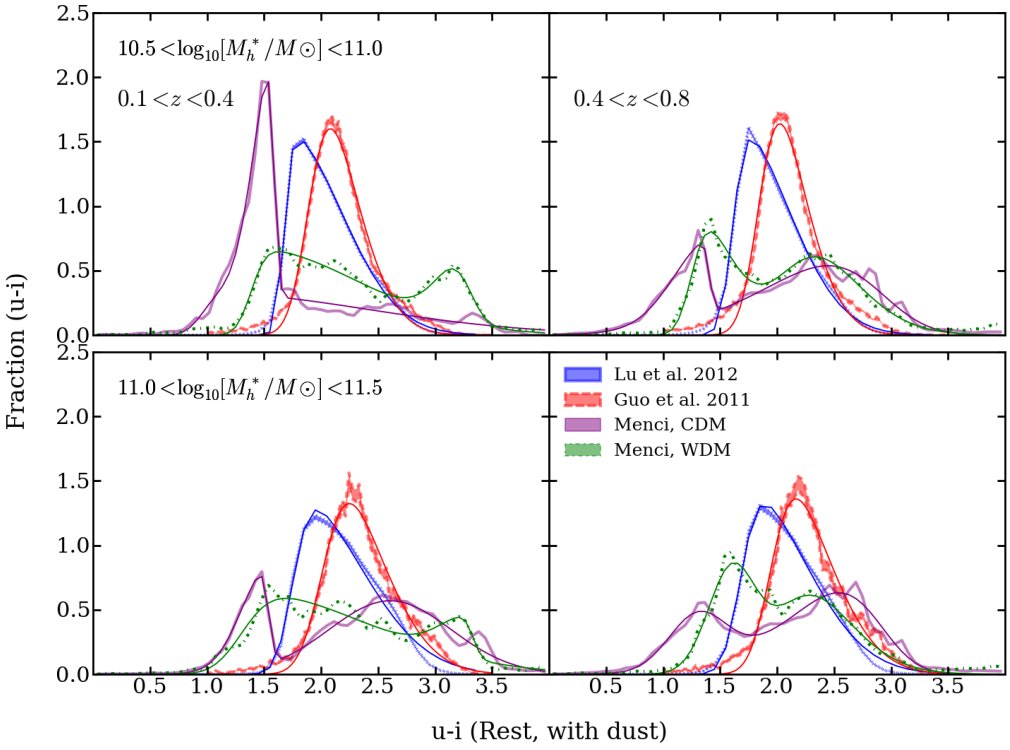

Figure 4.5: The normalized rest-frame distribution of colors, with dust extinction for the satellite galaxies plotted in Figure 2, along with the best fitting skewed normal distributions. The Menci models are fit with a sum of two skew normal distributions, to account for the prominent secondary peak that appears.

Host Properties	Dust	Guo et al. 2012 c_o, σ, a	Lu et al. 2012 c_o, σ, a	Menci, CDM c_o, σ, a, $A2/A1$, $c_{o,2}$, σ_2, a_2	Menci, WDM c_o, σ, a, $A2/A1$, $c_{o,2}$, σ_2,a_2
$10.5 < \log[M_*/M_\odot] < 11.0$, $0.1<z<0.4$	Yes	1.7, 0.5, 8.8	1.9, 0.4, 2.5	1.6, 0.2, -5.0 1.4, 1.0, 1.5, 12.5	1.4, 1.0, 9.3 0.2, 3.3, 0.25, -1.9
	No	2.0, 0.3, 1.5	1.7, 0.5, 7.0	1.3, 0.3, 1.4	1.9, 0.5, -4.0
$10.5 < \log[M_*/M_\odot] < 11.0$, $0.4<z<0.8$		1.6, 0.5, 5.6	1.8, 0.3, 2.4	1.4, 0.3, -9.0 3.1, 2.9, 0.9, 2.5	1.3, 0.3, 3.0 1.5, 2.1, 0.5, 1.6
		1.9, 0.3, 1.2	1.6, 0.4, 4.5	1.7, 0.5,-5.4	1.7, 0.4, -4.4
$11.0 < \log[M_*/M_\odot] < 11.5$, $0.1<z<0.4$		1.8, 0.6, 5.1	1.2, 0.5, 2.7	1.5, 0.24, -5.9 4, 2.6, 0.5, 0	1.3, 1.1, 5.5 0.3, 3.3, 0.3,-4.5
		2.5, 0.3, -1.2	1.8, 0.5, 4.3	2.0, 0.5,-4.4	1.9, 0.5, -4.2
$11.0 < \log[M_*/M_\odot] < 11.5$, $0.4<z<0.8$		1.7, 0.5, 4.6	1.9, 0.5, 2.9	1.3, 0.25, 0 2.3, 2.9, 0.7, -2.6	1.6, 0.2, 0 1.3, 2.0, 0.5, 1.7
		2.4, 0.3, -1.3	1.7, 0.5, 3.7	1.7, 0.5,-6.2	1.7, 0.4, -5.3

Table 4.1: Best fit parameters of the color distributions

some of these observations, a more detailed discussion of many of these comparisons can be found in the original papers describing the models (Menci et al. 2012; Guo et al. 2011b; Lu et al. 2012).

4.4.1 Satellite luminosity function

Studies of the satellite luminosity function typically focus on satellite galaxies at low redshifts with $z < 0.1$, both in simulation and observations. As we have shown in the previous chapter, the COSMOS field can be used to study satellites at higher redshifts, and yields luminosity functions at low redshifts which are consistent with satellite luminosity functions from SDSS (e.g. Guo et al. 2011a; Liu et al. 2011; Lares et al. 2011; Strigari & Wechsler 2012), and the Milky Way satellite luminosity function (Tollerud et al. 2008). All four of the models provide good fits to the satellite luminosity function of Milky Way mass hosts at low redshifts, by design. The extra information of redshift and host

galaxy stellar mass provides additional constraints to these models. For instance, when comparing to bright field galaxies observed by Pérez-González et al. (2008), Guo et al. (2011b) found their model agreed well at low redshift, while becoming discrepant with observation by a redshift of 1.

4.4.2 Satellite Colors

In Section 4.3.3 we showed that the three CDM models predicted significantly different distributions for the colors of faint satellite galaxies. Although we cannot directly test this prediction in this work, some comparisons can be made with measurements of field galaxies at low redshifts. In particular, Guo et al. (2011b) found that their model predicted colors matched observed SDSS colors well except for low masses ($\log_{10}[M^*/M_\odot] < 9.5$), which were redder in the simulation than in observations. Menci et al. (2012) found that down to an absolute magnitude of $M_r = -18$ (roughly a stellar mass of $\sim$9 $\log[M^*/M_\odot]$), their color distribution agreed well with SDSS measurements from Baldry et al. (2004). They did not compare for fainter satellites.

The Lu and Guo semi-analytic models were not optimized to reproduce the observed galaxy color-magnitude relation at low redshifts, thus future implementations of these semi-analytic models which include this information may yield significantly different predictions for the color distribution of satellite galaxies.

Recently, Knobel et al. (2012) measured the red fraction of massive ($\log_{10}[M^*/M_\odot] > 10$) satellites between redshift 0.1 and 0.8, and found no evidence for significant evolution over this time, indicating that these low redshift results may apply to higher redshifts.

From a different point of view, Behroozi et al. (2012) used abundance matching techniques to infer that low stellar mass ($\log_{10}[M^*/M_\odot] \sim 9$) *field* galaxies on average continue to form a significant fraction of their stars at redshift ~ 1. However, this result cannot be directly applied to satellite galaxies, as star formation in satellite galaxies has been shown to be quenched relative to field galaxies at low redshifts (Pasquali et al. 2010; Geha et al. 2012; Kauffmann et al. 2004).

Future measurements of faint satellite colors since intermediate redshifts will provide interesting new constraints on these models, by helping to distinguish between suppression of the satellite luminosity function by environmental quenching and supernovae feedback, or the subhalo mass function at the low mass end.

4.5 Discussion and conclusion

We have compared predictions from three current semi-analytic models applied to CDM simulations and one semi-analytic model applied to a WDM simulation to observations of satellite galaxies in two redshift intervals and with two host mass bins. By comparing multiple models, we demonstrated that the effects of varying star-formation prescriptions within semi-analytic models, and varying the underlying dark matter mass function have fundamentally different effects on the predicted host stellar mass dependence, and redshift evolution of the satellite luminosity function. In particular, the WDM model predicts a satellite luminosity function with much weaker host stellar mass dependence and redshift evolution than any of the CDM models.

We find that the WDM provides the best match to observation in all redshift and

host mass intervals, most closely matching the host mass dependence and lack of redshift evolution in the data. This comparison highlights the importance of comparing models of satellite galaxy evolution to observations from a range of redshifts and stellar mass regimes, as a model that provides a close match to observation in one regime can perform more poorly than other models in a different regime.

This exploration of different CDM models suggests that *current* models generically have difficulty reproducing the mass and redshift dependence of the satellite luminosity function. However, future improvements to the semi-analytic models for star formation prescriptions may allow for improvements between the CDM predictions of the satellite luminosity function. In fact, as shown in Chapter 3 subhalo abundance matching techniques (Busha et al. 2011) can reconcile our observations with the CDM subhalo mass function. Although abundance matching is descriptive and not directly linked to known physical processes, it is possible that future semi-analytic models will be able to produce the stellar mass to halo mass relation as illustrated by this approach.

Numerous studies have shown that abundance matching techniques can be used to map measured properties of star formation to simulated halo masses, and thus provide important constraints for semi-analytic models (e.g Behroozi et al. 2012; Reddick et al. 2012; Behroozi et al. 2010; Busha et al. 2011). The descriptive power of abundance matching is limited for low mass satellite galaxies by the resolution of dark matter simulations and the depth and redshift range of observations. Higher resolution simulations with varying dark matter power spectra, in conjunction with environmentally dependent abundance matching, can potentially provide very interesting constraints for the physical processes governing star formation in low mass halos in varying cosmologies.

Additional observational data are needed to further constrain the models and therefore help distinguish between whether the discrepancy between predicted and observed satellite luminosity function can be mitigated by either improved baryonic physics in CDM models or a WDM power spectrum. For example, we have shown in this work that the Lu CDM model which most closely matches the data predicts a significantly different color distribution for the satellite galaxies than the Menci WDM model, indicating that future observations of the colors of faint satellites will provide an important test of whether the luminosity function of satellites is suppressed primarily by baryonic processes or by a WDM mass function. We plan to carry out such measurement in the near future by exploiting the rich multicolor datasets publicly available in the HST legacy fields.

Acknowledgements

A.M.N. and T.T. acknowledge support from NSF through CAREER grant. T.T. acknowledges support from the Packard Foundations through a Packard Research Fellowship. We sincerely thank A. Benson, A. Peters, R. Wechsler and F. van den Bosch for useful comments and insightful conversations. This work was based on observations made with the NASA/ESA Hubble Space Telescope, and obtained from the Data Archive at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with the COSMOS and GOODS projects. Part of the work presented in this chapter was performed by AMN and TT while attending the program "First Galaxies and Faint Dwarfs: Clues to the Small Scale Structure of Cold Dark Mat-

ter" at the Kavli Institute of Theoretical Physics at the University of California Santa Barbara, supported in part by the National Science Foundation under Grant No. NSF PHY05-51164.

Chapter 5

The luminous satellites of gravitational lenses

This chapter was published as Nierenberg, A. M., Oldenburg, D., Treu, T., "Do gravitational lens galaxies have an excess of luminous substructure?", MNRAS, 436, 2120, (2011) and is included here with minor formatting adjustments

As was discussed in the introduction, outside of the Local Group, a powerful way to measure the subhalo mass function without relying on the presence of baryons is with strong gravitational lensing by a galaxy deflector. When a luminous background source is strongly lensed by an intervening galaxy, the number, positions, and magnifications of the images that appear depend only on the mass distribution of the lens galaxy and the relative angular positions of the lens and the source (see Treu 2010, and references therein). Substructure along the line of sight can be identified via a deviation in the magnification, positions or time delays of lensed images from what would be expected given a smooth deflector mass distribution (e.g. Mao & Schneider 1998; Metcalf & Madau 2001; Keeton & Moustakas 2009).

Gravitational lensing has been applied to a variety of lens systems in order to detect subhaloes and place constraints on their mass function (e.g. Dalal & Kochanek 2002; Vegetti et al. 2012; Fadely & Keeton 2012; Xu et al. 2009, 2013a). Due to the small

numbers of systems available for this type of analysis, the uncertainties on the inferred subhalo mass function remain large. However, future wide-area surveys such as LSST and PANSTARRS are expected to find thousands of new lens galaxies (Oguri & Marshall 2010) which can be analysed for substructure. Furthermore, the next generation of large telescopes and adaptive optics will make the deep, high resolution imaging required for the analyses of lensed images fast and therefore feasible for a large number of lenses, with sensitivity extending to lower masses.

In order to use the subhalo mass function measured around gravitational lenses as a test of ΛCDM, it is crucial to understand the lensing selection function, and in particular whether lens galaxies are representative of field galaxies. Several works have used simulations of different galaxy scale lenses with background point sources to determine which dark matter halo properties, and survey selection criteria affect the gravitational lensing cross section (Keeton & Zabludoff 2004; Mandelbaum et al. 2009; van de Ven et al. 2009; Arneson et al. 2012; Dobler et al. 2008) . These studies found that the lensing cross section is by far most strongly dependent on the surface mass density of the deflector.

Treu et al. (2009), showed that lens galaxies in the Sloan Lens Advanced Camera for Surveys (SLACS) (Bolton et al. 2004, 2005, 2006; Auger et al. 2009) exist in environments with densities which are consistent with those of non lens galaxies with similar stellar masses and velocity dispersions. Fassnacht et al. (2011) found a similar result for a sample of intermediate redshift lens galaxies. Furthermore, Auger et al. (2010) and Kochanek et al. (2000) showed that the fundamental plane for SLACS and CLASS (Myers et al. 2003; Browne et al. 2003) lens galaxies is consistent with that of non-lens, early-type galaxies.

In order to generalise the measurements of the subhalo population mass function of lens galaxies it is essential that we understand whether lens galaxies have substructure populations typical of their non-lens counterparts. Outside of the Local Group, the comparisons between the subhaloes of lens and non-lens galaxies are only possible for those subhalos which contain a significant population of stars. Those subhaloes which contain stars are expected to be the most massive, and therefore affect the lensing cross section most strongly. We expect that the properties of the subhalo population which significantly alter the probability of lensing should be evident in the subhalos containing luminous galaxies which are the most massive. Of course, with this sort of analysis we cannot test whether there are indirect selection effects connecting lens galaxies and the dark subhalo population, which do not directly influence the lensing cross-section.

As an important test, Jackson et al. (2010) counted the number of projected companions within 20 kpc apertures around lens galaxies in SLACS and CLASS, and compared them with the number of objects around early-type galaxies selected from Sloan and COSMOS ACS imaging. Jackson et al. (2010) found that SLACS lenses had a similar number of companion objects to other early-type galaxies in Sloan, while CLASS lenses had orders of magnitude more companions in projection than non-lens galaxies in COSMOS field galaxies.

In this work we revisit the question of the populations of luminous companions of gravitational lenses, taking into account several important developments with regards to the detection and study of satellites. First, we apply the image processing and statistical analysis developed in Chapters 2 and 3 to ACS images of SLACS lenses, which enables us to directly study the satellite population of these galaxies to more than a thousand

times fainter than the host galaxies, and very close in projection ($\sim$1.″0).

Secondly, we infer, for the first time, the dependence of the number of satellites on the host galaxy stellar mass for both SLACS lenses and COSMOS field galaxies. This allows us to rigorously incorporate differences in the distribution of host stellar masses when comparing between the two samples, which is essential given the strong observed dependence on the number of satellites as a function of host stellar mass as seen in Chapter 3.

The Chapter is organised as follows: in §5.1 we describe the selection of the sample of lens and non-lens host galaxies around which we study satellite galaxies. In §5.2 we describe our procedure for detecting companion objects around each of the samples of hosts. In §5.3 we explain the statistical method we use to infer the typical number of satellites per host galaxy as a function of stellar mass and redshift. In §5.4 we compare the inferred number of satellites per host for the SLACS and non-lens galaxies. In §5.5 we compare the number of satellites around COSMOS and CLASS hosts. Finally in §5.6 we conclude with a discussion and summary of the results.

5.1 Host Galaxy Selection

We restrict the galaxies analysed in this work to those which are imaged with sufficient depth and resolution to allow us to detect faint satellite galaxies. Furthermore, as the number of satellites is known to depend on the stellar mass of the host galaxy, we study only systems which had previously estimated stellar masses or multiple bands of photometry and a measured redshift, which enables us to estimate the stellar mass of

the systems. Figure 5.1 contains a summary of the redshift, stellar mass and apparent magnitude distributions of the host galaxies.

5.1.1 Lens Host Galaxies

In this work we focus mainly on measuring the properties of satellites around SLACS lens galaxies. SLACS candidates were detected in Sloan spectroscopy by searching for spectra which were composed of early-type, low redshift lenses superimposed with a higher redshift emission-line galaxy. Lens status was then confirmed using either ACS or WFPC2 imaging to confirm the morphological characteristics expected in gravitational lensing (Bolton et al. 2004, 2005, 2006).

We restrict our comparison to those confirmed SLACS early-type lenses with redshifts $z > 0.1$ which have deep (multiple exposure) I_{814} imaging in order to enable us to measure the satellite population using the same tools which we apply to COSMOS imaging. We also require that the lenses were imaged with ACS rather than WFPC2, in order to ensure that there was a sufficient area of deep, high-resolution imaging around the lenses to allow for an accurate measurement of the satellite population. We excluded one system with detector artefacts very near the lens due to a saturated star, and one system which had four companions of comparable brightness in the field, making the detection of faint satellites impractical. The final sample contained 32 hosts. Host stellar masses, redshifts and photometry were obtained from Auger et al. (2010).

5.1.2 Non-Lens Host Galaxies

We select non-lens host galaxies from the COSMOS ACS survey. We study host galaxies with stellar masses greater than $M^* > 10^{10.5} M_\odot$ as measured by Ilbert et al. (2010) using ground based photometry in conjunction with ACS imaging (Auger et al. 2010). Hosts are restricted be between redshifts $0.1<z< 1$ and to not be within R_{200} of a more massive companion, where R_{200} is estimated using the observed relationship between stellar mass and R_{200} from (Dutton et al. 2010). This isolation requirement is included to ensure that we are not studying host galaxies which are themselves satellites of a more massive system. We also use the strong gravitational lensing catalog by Faure et al. (2008) in order to exclude strong gravitational lenses from our sample.

COSMOS host stellar masses are based on SExtractor's MAGAUTO (Bertin & Arnouts 1996) from ACS imaging, unlike SLACS photometry which was performed using deVaucouleurs profile fitting. However, for bright galaxies, we found that this did not make a difference in COSMOS photometry (see also Benítez et al. 2004).

5.2 Companion Detection

We use the same method to detect companion objects around SLACS and COSMOS host galaxies as was used in Chapters 2 and 3.

For COSMOS host galaxies, we compare our detections to those detections already in the COSMOS ACS photometric catalogs. New detections are added, while we replace the photometry for objects which were already in the photometric catalogs within ~ 20 $r_{\rm eff}$ of the host galaxies, with our own, thereby ensuring that the host light does not

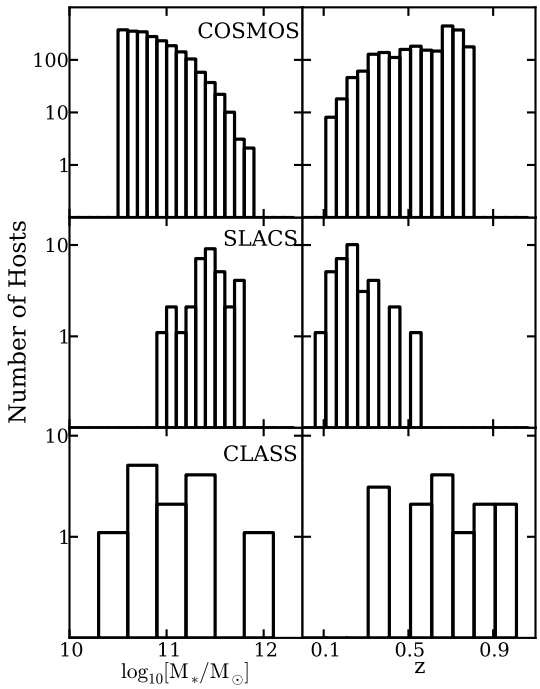

Figure 5.1: Stellar mass and redshift distributions of host galaxies considered in this work. Lens galaxies are systematically more massive than non-lens galaxies. This is expected given that the lensing cross section increases as velocity dispersion to the fourth power.

contaminate nearby objects.

For SLACS we detect objects in the full ACS images around the lens galaxies and then add in the new detections from the subtracted cutouts in the same way we do for COSMOS. For the SLACS images, bright stars and detector artifacts were masked by hand.

For both CLASS and SLACS hosts, we study all objects with $I_{814} < 25$ magnitudes AB.

5.3 Inference of the Number of Satellites Per Host

In order to study satellite galaxies using single band photometry, we use the statistical method discussed in Chapter 2, with two new additions to the model. These additions allow for an accurate comparison between samples of host galaxies with different stellar mass distributions, and uncertain virial radii. We conclude by describing the priors used for the model. The likelihood function and a more detailed description of the statistics, including how we adjust for masked regions in the images can be found in Appendix 5.7.

5.3.1 Number of satellites as a function of host stellar mass

In order to account for the differences in the stellar mass distributions of COSMOS and SLACS host galaxies (Figure 5.1), we infer the number of satellites as a continuous function of host stellar mass for the two samples.

$$N_s \propto N_{s,o}(1 + \log_{10}[M^*_{host}/M\odot] - 11.4)^{\kappa_s} \tag{5.1}$$

Given that the amount of correlated structure is also a function of the host stellar

mass, we also infer the number density of background objects as a function of stellar mass:

$$\Sigma_b = (\Sigma_{b,o} + \kappa_b(\log_{10}[\mathrm{M}^*_{\mathrm{host}}/\mathrm{M}\odot] - 11.4))\,(1+z)^{\delta_{\mathrm{z,b}}} \tag{5.2}$$

This redshift evolution term in the foreground/background number density is not necessary for SLACS which covers a much smaller redshift range than the COSMOS hosts. We do not study redshift evolution in the number of satellites per host as there was no strong dependence apparent in Chapter 3.

5.3.2 Uncertainty in the virial radius

In order to account for variations in host halo mass and projected angular sizes due to redshift variations, it is important to choose an appropriate distance scale within which to study the satellite population. In 3 we scaled all distances by R_{200} as estimated by the observed relationship between stellar and halo mass from Dutton et al. (2010). This relationship becomes increasingly uncertain for high stellar masses (e.g. Behroozi et al. 2010).

The median stellar mass of the SLACS lens galaxies is $\log[\mathrm{M}^*_{\mathrm{host}}/\mathrm{M}\odot] = 11.4$, which is approximately where the halo mass to stellar mass relation becomes very uncertain (e.g. Behroozi et al. 2010; Leauthaud et al. 2012). To account for this, for all hosts with stellar masses greater than $\log[\mathrm{M}^*_{\mathrm{host}}/\mathrm{M}\odot] > 11.4$, the true virial radius is included as a model parameter, with mean and standard deviation given by Dutton et al. (2010). We further restrict the inference to only allow for virial masses $\mathrm{M}_{\mathrm{vir}} < 10^{14}\mathrm{M}\odot$, as there are no clusters in either the COSMOS or SLACS samples, given group richnesses and lensing

profiles. For our final results, we marginalise over the inferred virial radii. We note that the exact threshold where virial mass uncertainty is included does not significantly affect our results, as at lower stellar masses the stellar mass to virial mass uncertainty becomes smaller and the uncertainty due to the significantly smaller number of satellites per host dominates.

5.3.3 Priors

We use Gaussian priors on the parameters describing the number density of background/foreground objects ($\Sigma_{\rm b,o}$, κ_b and $\delta_{\rm z,b}$). We obtain these priors by measuring the background properties in annuli between $1.0 < r < 1.5$ $\rm R_{200}$ around the COSMOS hosts. For SLACS, we use priors with the same means as in COSMOS, but with broader standard deviations to allow for the fact that the COSMOS field is on average denser than pure-parallel fields (Fassnacht et al. 2011). We do not allow for redshift variation in the background/foreground number density in SLACS due to the small sample size and narrow redshift range of the central galaxies.

For COSMOS satellites, we apply a prior on the slope of the projected radial profile of $\gamma_{\rm p}$ from 3. Values for mean and standard deviations for these priors are listed in Table 5.1. We infer the properties of SLACS satellites with and without this prior on $\gamma_{\rm p}$.

We infer all parameters in bins of magnitude offset between companion and host galaxy ($\Delta \rm m = m_{sat} - m_{host}$).

Parameter	Description	COSMOS Prior	SLACS Prior
$N_{s,o}$	Number of satellites for hosts with $M^* = 10^{11.4} M_\odot$	U(0,100)[a]	U(0,100)
κ_s	Dependence of satellite number on host stellar mass	U(0,60)	U(0,60)
γ_p	Projected radial slope of the satellite number density	N(-1.1,0.3)[b]	U(-9,0)
$\Sigma_{b,o}$	Number of background/foreground (b/f) objects per arcmin^2 with $I_{814} < 25$	N(47,0.3)	N(47,2)
α_b	Slope of the b/f number counts as a function of magnitude	N(0.300,0.005)	N(0.3,0.1)
κ_b	Dependence of the number of b/f objects on host stellar mass	N(1.5,0.3)	N(1.5,0.3)
$\delta_{z,b}$	Dependence of the number of b/f objects with redshift.	N(0.03,0.02)	...

Table 5.1: Priors used in the inference of the number and spatial distribution of satellite galaxies around SLACS and COSMOS host galaxies. (a) Uniform prior between (min,max), (b) Normal prior with (mean, standard deviation)

Δ m	$N_{s,o}$	κ_s	γ_p	$\Sigma_{b,o}$	α_b	$N_{s,o}$ [a]	κ_s [a]
4	6^{+3}_{-3}	$2.1^{+1}_{-0.7}$	$-0.7^{+0.2}_{-0.3}$	47^{+2}_{-2}	$0.38^{+0.05}_{-0.04}$	3^{+2}_{-2}	3^{+1}_{-1}
5	9^{+5}_{-3}	$2.4^{+1}_{-0.8}$	$-0.8^{+0.2}_{-0.3}$	47^{+2}_{-2}	$0.36^{+0.04}_{-0.02}$	7^{+3}_{-3}	$2.8^{+1}_{-0.9}$
6	20^{+5}_{-4}	$2.7^{+0.5}_{-0.5}$	$-0.7^{+0.1}_{-0.1}$	47^{+2}_{-2}	$0.39^{+0.02}_{-0.02}$	18^{+4}_{-4}	$2.7^{+0.7}_{-0.6}$
7	28^{+7}_{-6}	$2.7^{+0.5}_{-0.5}$	$-1.0^{+0.1}_{-0.2}$	46^{+2}_{-2}	$0.39^{+0.02}_{-0.02}$	27^{+5}_{-5}	$2.7^{+0.5}_{-0.5}$
8	37^{+9}_{-7}	$3.4^{+0.7}_{-0.5}$	$-1.0^{+0.1}_{-0.2}$	42^{+2}_{-2}	$0.39^{+0.02}_{-0.02}$	38^{+7}_{-7}	$3.5^{+0.6}_{-0.6}$

Table 5.2: Summary of inference results for SLACS satellites, see Equations 5.1 and 5.2. Inferred between $0.03 < R < 0.5\ R_{200}$. (a) Inferred with a Gaussian prior on γ_p with mean -1.1 and standard deviation 0.3

5.4 Results

In Tables 5.2 and 5.3, we present the results of our inference on the cumulative number of satellites per host galaxy as a function of stellar mass and host redshift, in two bins of magnitude offset between host and companion objects $\Delta m = m_{sat} - m_{host}$.

The results can be summarized as follows. For both COSMOS and SLACS hosts, we detect a significant dependence of the number of satellites per host on the host stellar mass, which is consistent between the samples. These results are also consistent with the results of Chapter 3, in which we observed that the number of satellites per host at fixed Δ m was significantly higher for hosts with $11 < \log[M^*_{host}/M_\odot] < 11.5$ than for hosts

Δ m	$N_{s,o}$	κ_s	γ_p
2	$0.8^{+0.2}_{-0.2}$	$1.4^{+0.3}_{-0.3}$	$-1.1^{+0.1}_{-0.1}$
3	$1.7^{+0.3}_{-0.3}$	$1.7^{+0.3}_{-0.3}$	$-1.2^{+0.1}_{-0.1}$
4	$2.9^{+0.4}_{-0.4}$	$1.6^{+0.3}_{-0.2}$	$-1.1^{+0.1}_{-0.1}$
5	$4.2^{+0.7}_{-0.8}$	$1.8^{+0.4}_{-0.3}$	$-1.1^{+0.1}_{-0.1}$
6	12^{+1}_{-1}	$2.3^{+0.4}_{-0.3}$	$-0.6^{+0.1}_{-0.1}$
7	20^{+3}_{-3}	$1.8^{+0.5}_{-0.5}$	$-0.8^{+0.2}_{-0.2}$

Table 5.3: Summary of inference results for COSMOS satellites inferred between $0.03< r < 0.5\ R_{200}$, see Equations 5.1 and 5.2.

with $10.5 < \log[M^*_{host}/M_\odot] < 11$.

We find that the slope of the power-law radial profile of the satellite number density of satellites around SLACS lenses is consistent with that of COSMOS field galaxies and goes approximately as $\gamma_p \sim -1$. As discussed in Chapters 2 and 3, this slope is approximately isothermal. Given the range of host concentrations across our sample, and the fact that we are fitting only a single power-law to the radial profile of satellites, this result can be reproduced by a sum of Navarro Frank and White radial profiles (Navarro et al. 1996) for the satellite number density, which is approximately the density profile which would be expected if satellite galaxies followed the density profile of the host halo, as predicted by simulations (e.g Kravtsov 2010).

Therefore, we perform the inference a second time, for SLACS galaxies using the same Gaussian prior on the projected radial profile as was used for COSMOS. These results are listed in the last two columns of Table 5.2. The Gaussian prior lowers the uncertainty in the inferred number of satellites, but it does not significantly affect the number of satellites per host.

In Figure 5.2, we compare the inferred number of satellites for COSMOS and SLACS

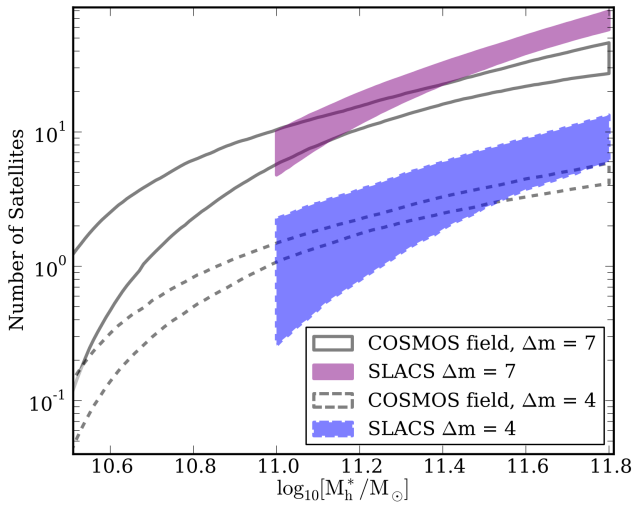

Figure 5.2: The average cumulative number of satellites per host between 0.03< R_{200} <0.5, near COSMOS and SLACS hosts, for two values of Δm where all objects analysed are restricted to have $m_{\rm obj} - m_{\rm host} < \Delta$m . The filled curves show the one sigma confidence intervals for the inferred number of satellites per host as a continuous function of the host stellar mass. The inferred values for these functions are listed in Table 5.3. These results are inferred with a Gaussian prior on the projected slope of the satellite radial profile $\gamma_{\rm p}$ with mean -1.1 and standard deviation 0.3.

hosts using the same prior on $\gamma_{\rm p}$. The number of satellites as a function of host stellar mass is consistent at all stellar mass values, although there may be a small trend in a steeper slope for the SLACS satellites, the difference is significant only at the $\sim 1\sigma$ level. Note that although values of individual parameters are not always consistent for COSMOS and SLACS, the average number of satellites as a function of host stellar mass is, due to degeneracies between the parameters κ_s and $\rm N_{s,o}$, and the final number of satellites as a function of stellar mass. In Figure 5.3 we show the bivariate posterior probability distributions for κ_s and $\rm N_{s,o}$ for Δm = 4 and 7.

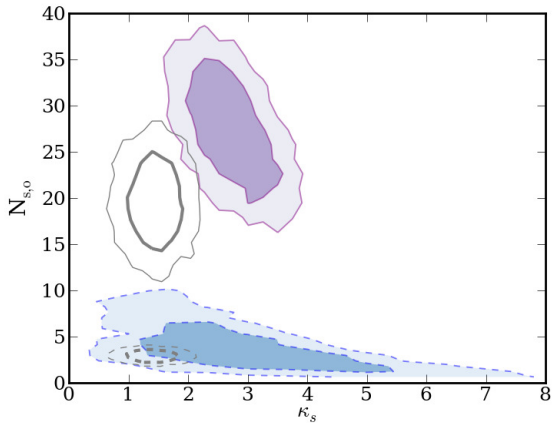

Figure 5.3: Comparison between posterior probabilities for the parameters κ_s and $N_{s,o}$ for Δm= 4 and 7, for SLACS and COSMOS host galaxies. The colours have the same meaning as in Figure 5.2. The darker inner contours and the lighter outer contours represent the 68 and 95 % confidence intervals respectively.

5.5 Comparison with CLASS

Although the focus of this work is with the satellite population of SLACS lenses, it is interesting to see whether the CLASS lenses have a relatively higher rate of companion detection than COSMOS galaxies, given the stellar mass distribution of CLASS hosts, and the results for our inference on the dependence between satellite number and host stellar mass.

Gravitational lenses in the Cosmic Lens All Sky Survey (CLASS, Myers et al. 2003; Browne et al. 2003) were selected by searching for flat-spectrum radio sources with multiple spatial components. This selection is very distinct from the SLACS selection, as it is purely based on the detection of a gravitationally lensed radio source, and has no requirements for the detection of the deflector galaxy. For this reason, the CLASS galaxies span a much larger range in redshift space than SLACS galaxies, and they do not all have well measured photometry or redshifts. For this comparison we consider only

CLASS hosts which have I_{814} imaging with either WFPC2 or ACS [1] and spectroscopic redshift measurements.

We obtained stellar masses of the CLASS lenses by compiling photometry from the literature (Lagattuta et al. 2010; Rusin & Kochanek 2005; Muñoz et al. 1998), and applying the stellar population synthesis modelling code by Auger et al. (2010). Due to the non uniform methods of obtaining photometry across the different works, we assume a factor of ~ 2 uncertainty in estimated stellar masses. We further restrict our sample to those lenses which have estimated stellar masses greater than $M^*/M_\odot > 10^{10.4}$ in order to allow us to continue to use the functional form from Equation 5.1. The final comparison sample consists of 14 hosts [2] with properties shown in Figure 5.1.

We use companion object detection from Jackson et al. (2010) who used SExtractor to detect objects within 20 kpc of the CLASS deflectors, with an I_{814} magnitude limit of 24.9. Jackson et al. (2010) visually excluded lensed images to ensure they were not counted as companion objects. After these considerations, Jackson et al. (2010) detects four companion objects within 20 kpc of the 14 CLASS host galaxies which have spectroscopic redshift measurements, and meet our stellar mass requirements.

Using the stellar mass distribution of the CLASS hosts, we apply the results from the previous section to estimate the number of companion objects we would have expected around COSMOS hosts with the same stellar mass distribution. To do this, we drew randomly ten thousand times from the posterior probability distribution functions for κ_s,$\gamma_{\rm p}$ and $N_{\rm s,o}$, as well as from Gaussian distributions centred on the estimated stellar

[1] with the exception of 1422 which has I_{791} imaging

[2] CLASS B0128+437, MG0414+0534, CLASS B0445+123, CLASS B0631+519, CLASS B0712+472, CLASS B1030+074, CLASS B1152+199, CLASS B1422+231, CLASS B1608+656, CLASS B1933+503, CLASS B1938+666, CLASS B2045+265, CLASS B2108+213, CLASS B2319+051

masses with standard deviations of 0.3 dex in order to account for uncertainty in the host stellar mass estimates. We used the values of $\gamma_{\rm p}$ to rescale the expected number of satellites to the smaller area studied by Jackson et al. (2010).

Assuming that the closest satellites could be found only as close as 10 kpc, to the host centres, due to obscuration by the host galaxy, a sample of 14 COSMOS hosts with the same stellar masses, and brightnesses as the CLASS hosts would have about 4 ± 2 satellites, and approximately 4 ± 2 background/foreground objects, depending on the assumed background/foreground number density around CLASS hosts. Given the small number statistics, this prediction is marginally consistent with the detection of four objects around CLASS hosts, however these numbers do not take into account obscuration by lensed images.

We can roughly account for obscuration by lensed images, by assuming that the lensed images are found in annuli at the lens Einstein radii, with radial width of $0\farcs4$. For double image lenses, we assume that about one third of the annulus is blocked by lensed images, and for four image lenses, we assume the full annulus is blocked, although the exact fraction of blocking does not significantly affect the results. Using this approximate accounting for lensed image obscuration, we expect 2-3 satellite galaxies and 1-2 foreground/background objects to be detected around CLASS host galaxies.

Although the main search radius in Jackson et al. (2010) was within 20 kpc, many of the companion objects around CLASS lenses were found even closer to the host galaxy, and brighter than the detection limit of 24.9. If we instead consider the number of objects between 4 and 14 kpc, within 3.3 magnitudes of the central galaxy magnitude, our model would predict about 2 satellites and 1 background/foreground object, including the effects

of obscuration which are less relevant this close to the central galaxy, in comparison with the four companion objects Jackson et al. (2010) detected which met these criteria.

Therefore, taking into account the fact that the CLASS host galaxies are very massive, we find that they have approximately the same number of nearby companion objects as we observe around COSMOS host galaxies with comparable stellar masses. This result highlights the importance of considering host stellar masses when comparing samples of satellite galaxies.

5.6 Summary

We have measured the spatial distribution and number of satellites per host as a function of host stellar mass for SLACS lens galaxies and COSMOS field galaxies using host light subtraction, object detection and an updated version of the statistical analysis developed by Nierenberg et al. (2011, 2012). Furthermore, we compared the number of projected companion objects found within 20 kpc apertures of CLASS lens galaxies by Jackson et al. (2010) to what we would have expected around COSMOS field hosts with similar stellar masses

Our main results are summarised below:

1. We detect a significant population of luminous satellites around SLACS lens galaxies. Parametrising the spatial distribution of satellites as $P_{\rm sat}(r) \propto r^{\gamma_{\rm p}}$ we find $\gamma_{\rm p} \sim -0.8 \pm 0.2$, which is consistent with the spatial distribution of luminous satellites around COSMOS non-lens galaxies.

2. Parametrising the number of satellites per host within a fixed magnitude offset

from the host galaxy ($\Delta m = m_{sat} - m_{host}$) to be $N_s \propto N_{s,o}(1 + \log_{10}[M^*_{host}/M_\odot] - 11.4)^{\kappa_s}$, we find κ_s to be $\sim 2.9 \pm 0.6$ for SLACS satellites as measured for hosts between $10^{11} < \log_{10}[M^*_{host}/M_\odot] < 10^{11.8}$, and $\sim 1.8 \pm 0.5$ for COSMOS satellites as measured between $10^{10.5} < \log_{10}[M^*_{host}/M_\odot] < 10^{11.8}$. Given degeneracies between κ_s and $N_{s,o}$, the number of satellites per host as a function of host stellar mass is consistent within the measurement uncertainties for all values of Δm.

3. Using the above results, we find that the number of close companions to CLASS lenses found by Jackson et al. (2010) is consistent with what we observe around COSMOS non-lens galaxies, taking the distribution of CLASS host stellar masses into account, as well as obscuration due to bright lensed images.

From these results we conclude that the subhalo mass function measured from these strong gravitational lenses is representative of the global subhalo mass function for haloes of the same masses.

Acknowledgements

A.M.N., D.O. and T.T. acknowledge support from the Packard Foundations through a Packard Research Fellowship. A.M.N. and D.O. also thank the Worster family for their research fellowship. We sincerely thank A. Sonnenfeld for estimating stellar masses of CLASS lenses, and B. Kelly, C. Fassnacht, and N. Jackson, and the anonymous referee for extremely helpful comments and discussions. This work was based on observations made with the NASA/ESA Hubble Space Telescope, and obtained from the Data Archive at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations

are associated with the COSMOS and GOODS projects.

5.7 Appendix: Likelihood function

Here we describe how we infer the model parameters $\boldsymbol{\theta}$ (listed in Table 5.1), which describe the number and spatial distribution of satellite and background/foreground galaxies between 0.03 and 0.5 R_{200} as a function of host stellar mass, within a fixed magnitude offset from the host galaxy ($m_{obj} - m_{host} < \Delta m$), given our data $\boldsymbol{D}$.

Using Bayes' theorem, the probability of the model parameters being true given the data is proportional to:

$$\Pr(\boldsymbol{\theta}|\boldsymbol{D}) \propto \Pr(\boldsymbol{D}|\boldsymbol{\theta})\Pr(\boldsymbol{\theta}) \tag{5.3}$$

Where $\Pr(\boldsymbol{\theta})$ is the prior on the model parameters.

For the j^{th} host in our sample, the data consist of a set of object radii relative to the host galaxy $\boldsymbol{r}_j$, the total number of objects around the host galaxy, $N_j^{\rm obs}$, and the host stellar mass $M^*_{{\rm h},j}$. The likelihood function can be decomposed into a product over each host galaxy of the individual likelihoods :

$$\Pr(\boldsymbol{\theta}|\boldsymbol{D}) \propto \prod_j^{N_{\rm h}} \Pr(\boldsymbol{d}_j|\boldsymbol{\theta}, M^*_{{\rm h},j})\Pr(\boldsymbol{\theta}) \tag{5.4}$$

For each host galaxy, the likelihood is the probability of measuring $N_j^{\rm obs}$ times the product of the likelihoods for each object position given the model parameters.

$$\Pr(\boldsymbol{d}_j|\boldsymbol{\theta}, \boldsymbol{h}_j) = \Pr(N_j^{\rm obs}|\boldsymbol{\theta}) \prod_i \Pr(r_i|\boldsymbol{\theta}, \boldsymbol{h}_j) \tag{5.5}$$

The first term on the right hand side is a Poisson probability comparing the model predicted number of objects with the observed number of objects

The model prediction for the number of objects is given by the sum of the model prediction for the number of satellite galaxies and the number of background/foreground objects, where the number of background/foreground objects is given by:

$$N_b = A\Sigma_b \tag{5.6}$$

Where Σ_b is defined in Equation 5.1, and A is the area (in arcminutes) in which objects were observable for a given host.

The model prediction for the observable number of satellites $N'_{s,j}$, for the j^{th} host is given by:

$$N'_{s,j} = N_{s,j} \frac{\int_{r_{\rm min,j}}^{r_{\rm max,j}} r^{\gamma_{\rm P}} f_j(r) r dr}{\int_{0.03}^{0.5} r^{\gamma_{\rm P}} r dr} \tag{5.7}$$

Where $N_{s,j}$ is the model prediction for the number of satellites per host between 0.03 and 0.5 $R_{200,j}$, as defined in Equation , $f_j(r)$ is the fraction of observable area as a function of radius, and $r_{\rm min,j}$ and $r_{\rm max,j}$ are the minimum and maximum observable radii respectively in units of $R_{200,j}$. Note that this is where uncertainty in the virial radius enters in to the likelihood function. For hosts with stellar masses greater than $10^{11.4}$ $M_\odot$, the virial radius is an additional model parameter.

The second term on the right hand side of Equation 5.5 is the likelihood of the observed objects being at their observed positions given the model parameters. As we do not know whether a certain object is a satellite or a foreground/background object, the total probability of it appearing at a given position is the sum of the probability of it being there if it was a satellite, and the probability of it being there if it was a

background/foreground object, weighted by the relative probabilities of an object being a satellite or a background/foreground object given the model parameters:

$$\begin{aligned}\Pr(r_i|\boldsymbol{\theta},\boldsymbol{h}_j) = \quad & \Pr(r_i|\boldsymbol{\theta},\boldsymbol{h}_j,S)\Pr(S|\boldsymbol{\theta},\boldsymbol{h}_j) \\ & +\Pr(r_i|\boldsymbol{\theta},\boldsymbol{h}_j,B)\Pr(B|\boldsymbol{\theta},\boldsymbol{h}_j)\end{aligned} \tag{5.8}$$

The terms $\Pr(S|\boldsymbol{\theta},\boldsymbol{h}_j)$, and $\Pr(B|\boldsymbol{\theta},\boldsymbol{h}_j)$ are the relative probabilities of something being a satellite or a background/foreground object given the model parameters. For example:

$$\Pr(S|\boldsymbol{\theta},\boldsymbol{h}_j) = \frac{N'_{s,j}}{N'_{s,j} + N_{b,j}} \tag{5.9}$$

The term $\Pr(B|\boldsymbol{\theta},\boldsymbol{h}_j)$ is defined analogously.

The probability of observing a satellite at position r_i is given by:

$$\Pr(r_i|\boldsymbol{\theta},\boldsymbol{h}_j,S) = \frac{r_i^{\gamma_\mathrm{p}+1} f_j(r)}{\int_{r_{\mathrm{min,j}}}^{r_{\mathrm{max,j}}} r^{\gamma_\mathrm{p}} f_j(r) r dr} \tag{5.10}$$

The probability of observing a background/foreground object at position r_i is:

$$\Pr(r_i|\boldsymbol{\theta},\boldsymbol{h}_j,B) = \frac{r_i f_j(r)}{\int_{r_{\mathrm{min,j}}}^{r_{\mathrm{max,j}}} f_j(r) r dr} \tag{5.11}$$

Chapter 6

Detection of dark matter halos with strong lensing of narrow-line emission

This chapter has been accepted for publication in MNRAS as Nierenberg, A. M., Oldenburg, D., Treu, T., "Detection of a substructure with adaptive optics integral field spectroscopy of the gravitational lens B1422+231", MNRAS, (2014) and is included here with minor formatting adjustments

As we discussed in the Introduction, the measurement of the subhalo mass function with gravitational lensing is currently limited by the small number of suitable systems. In this chapter we demonstrate that subhalos can be detected via perturbations to image magnifications using strongly lensed *narrow-line* quasar emission, as originally proposed by Moustakas & Metcalf (2003), by combining adaptive optics with integral field spectroscopy in order to obtain spatially resolved spectra of individual lensed images.

Strongly lensed *narrow-line* quasar emission provides an alternative means of measuring the subhalo mass function (Moustakas & Metcalf 2003) Quasar narrow-line emission at low redshift is observed to be typically extended over tens to hundreds of parsecs depending on the source luminosity (Bennert et al. 2002), which corresponds to milliarcseconds for typical source redshifts, well above the microlensing scale. A benefit to

studying narrow-line emission is that many more quasars have detectable narrow-line emission as opposed to radio emission. This makes this method ideal for measuring the subhalo mass function in the thousands of quad quasar lenses which are expected to be discovered based on their optical properties in ongoing and future surveys including PANSTARRS, DES and LSST (Oguri & Marshall 2010).

While promising, spatially resolved spectroscopy of narrow line emission of quasar lenses has proven difficult to attain. Sluse et al. (2012b) and Guerras et al. (2013) have measured lensed image spectra for pairs of quasar images, and for wide image separation double lenses, and observed differential lensing between the broad and continuum emission, which is evidence for microlensing. In order to study the lensing signal of substructure as discussed before, it is necessary to measure the spectra of images individually. Metcalf et al. (2004) and Keeton et al. (2006), achieved sufficient spatial resolution in order to measure the broad emission in individual quad quasar images, but neither of these studies detected narrow emission with sufficient signal to noise ratio.

In a pioneering work, Sugai et al. (2007) used integral field unit (IFU) optical spectroscopy in order to measure narrow [OIII] fluxes in the gravitational lens RXJ1131. The relatively wide separation between the lensed images ($\sim 1\farcs0$) made it possible to spatially resolve the lensed images, and place an upper limit on the mass of perturbing substructure. Interestingly, they found that unlike the continuum and broad fluxes, the [OIII] image fluxes did not deviate significantly from a smooth model prediction without substructure. In order to apply this method to a larger sample of systems and to probe lenses with configurations with smaller image separations (such as fold lenses) it is essential to obtain higher spatial resolution via HST (Keeton et al. 2006), or adaptive optics from the

ground. In order to extend this method to the larger population of quasar lenses which are typically at redshift 1 and above, it is essential to obtain higher spatial resolution in order to resolve typical image separations of $0\farcs5$ or less. In this work, we demonstrate for the first time that subhalos can be detected using strongly lensed narrow-line quasar emission of a high redshift (z=3.6) quasar by combining adaptive optics with integral field spectroscopy in order to obtain spatially resolved spectra of individual lensed images.

In this chapter, we present a measurement of narrow-line lensing with unprecedented accuracy using the OH Suppressing Infra-Red Imaging Spectrograph (OSIRIS Larkin et al. 2006) with laser guide star adaptive optics (LGS-AO) (Wizinowich et al. 2006) at Keck. We choose as initial case to study the famous system B1422+231 (Patnaik et al. 1992). Previous extensive studies across many bands of photometry (including radio) enable us to compare our measurement with more traditional methods.

The structure of the chapter is as follows: In §6.1 we present our observations. In §6.2 we describe our Bayesian method for combining multiple exposures to get sub-pixel sampling of the images and extract spectra optimally. In §6.3 we present the extracted spectra and show fits to the broad and narrow line components in each of the lensed images, and the inferred fluxes from each. In §6.4 we discuss the gravitational lens models we use to fit the observed image fluxes and positions. In §6.5 we present results from a gravitational lens model assuming a single perturbing subhalo. In §6.6 we discuss future prospects for this method. Finally in §6.7, we summarise our results. Throughout this chapter, we assume a flat ΛCDM cosmology with $h = 0.7$ and $\Omega_{\rm m} = 0.3$. All magnitudes are given in the AB system (Oke 1974) unless otherwise stated.

6.1 Observations

We observed the gravitationally lensed quasar B1422+231 on June 10 2012 at the W. M. Keck 10 m telescopes, using OSIRIS with LGS-AO corrections. To maximise the signal-to noise ratio we used the largest available pixel scale which is $0\farcs1$. We used broad-band K (λ =2.17 microns) filter which provided a spatial field of view of $\sim 1\farcs6$x$6\farcs4$ and a spectral range of 416 nm in the observed frame. At the time of our observations, the instrument had recently been moved and the full calibration had not yet been performed. As a result we were limited to a bigger wavelength range and smaller field of view that made it possible to only fit three of the four images in the field of view. We observed the system using 300 s exposures, dithering along the long axis of the field of view, with sub-pixel offsets in order to recover more spatial information, for a total of 2700s of integration. We used the OSIRIS data reduction pipeline to turn the raw CCD images from each exposure into rectified, telluric corrected, and sky emission-subtracted data 'cubes' in which the spatial information of the images appeared in the x-y plane of the images, and the wavelength varied along the third dimension.

6.2 Data Reduction

Our goal is to measure integrated line fluxes separately for each lensed image, taking into account all uncertainties regarding image positions, image deblending, and PSF properties. We do this in a three step process.

First, we infer the properties of the PSF and image positions in 'white' images, which we created by taking the variance-weighted mean of the data cube over the wavelength

region dominated by the broad and narrow line emission between rest frame 4700 - 5100 Å. These white images are effectively narrow band images of the system. Figure 6.1 shows the white image for a single exposure.

Next, we use a Bayesian inference in order to model the system as three point sources (e.g., the three observed lensed-QSO images). The centre of lens galaxy is not within the field of view (Figure 6.1), and is $\sim$ 3 magnitudes fainter than the QSO images, thus we do not include its contribution to light in the image. The model contains a set of global parameters that are the same for all exposures (the offsets between the three sources and their relative fluxes) and a set of parameters that vary from exposure to exposure (the sky level, the absolute location of the images, the total flux of the three images, and properties describing the shape of the PSF for each exposure). The PSF is modelled as two concentric Gaussians to represent the seeing halo and the AO corrected diffraction core. The inner diffraction core is fixed to be circular and to have a full width at half maximum (FWHM) of $0\farcs065$, based on the size of the telescope and the wavelength of observation (see also van Dam et al. 2006). The seeing halo Gaussian FWHM, ellipticity, position angle, and amplitude relative to the diffraction core Gaussian are free parameters. We run an Markov Chain Monte Carlo (MCMC) sampler for at least ten thousand steps in order to explore the full posterior probability distribution function, including degeneracies between the parameters. We tested this method by simulating images with the same image separations and signal-to-noise ratio as our data set, and found that we could recover the image positions to an accuracy of better than 5 mas, and the image fluxes to percent level accuracy.

The resulting measured image positions are given in Table 6.1. The relative image

positions positions are consistent with radio and HST measurements of the system. Our observations occurred during excellent conditions, with typical Strehl $\sim$ 0.2-0.3, and seeing of $\sim 0\farcs3$

Finally, we extract the spectra using the inferred PSF properties and image positions, relying on the fact that the PSF does not vary significantly over the short wavelength range we are considering ($\sim$ 180 nm observed). We draw sets of parameters describing the PSF and image positions from the MCMC chains obtained in the previous step. Then, for each set of parameters, for each wavelength slice in the data cube, we perform a χ-squared optimisation via linear inversion in order to find the best fit amplitudes for the lensed images simultaneously for all exposures, allowing for variations in the background flux as a function of wavelength and exposure. The final extracted spectrum for each image is then given by the mean and standard deviation of the best fit image fluxes in each wavelength slice after iterating over all of the parameter draws. This method allows us to quickly and robustly incorporate uncertainties in our inference of the image positions and PSF parameters directly into our spectral extraction. Image spectra are plotted in Figure 6.1.

6.3 Integrated Line Fluxes

Having extracted spectra for each of the lensed images, the next step in our analysis is to measure the integrated line fluxes by modelling the spectra as a sum of narrow H-β and [OIII], broad H-β and continuum emission[1].

[1]We also tested for the presence of broad iron nuclear emission, but did not find a significant contribution, and thus omitted it from the final analysis.

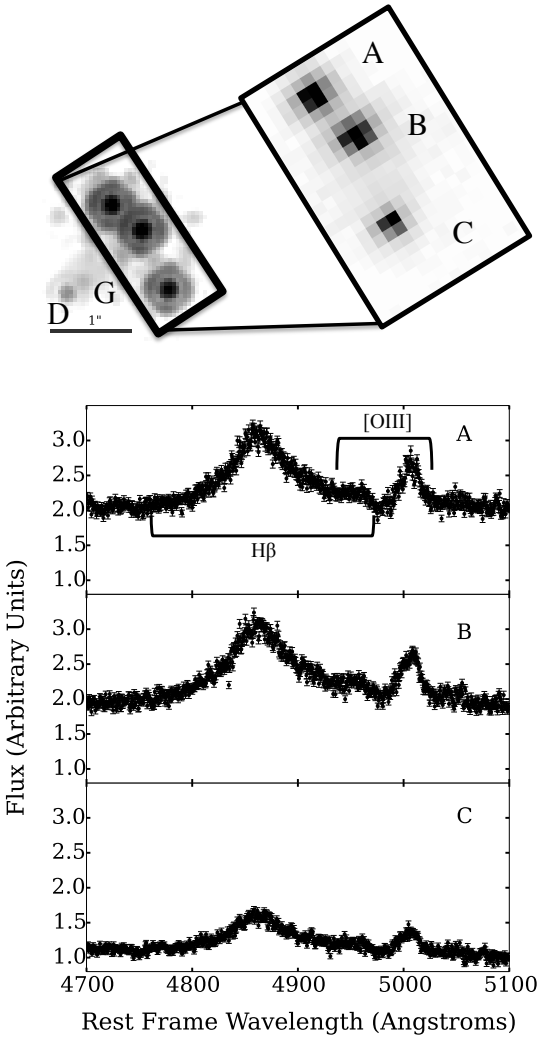

Figure 6.1: Upper Left: HST NICMOS F160W image. Upper Right: Single 300s K band exposure of target using OSIRIS with adaptive optics at Keck, we were restricted to the small field of view due to instrument difficulties. Lower panels: The final extracted spectra for each of the images.

Component	dRa	dDec	[OIII] flux	broad H-β	continuum
A	0.387 ±0.005	0.315 ±0.005	0.88 ± 0.01	4.67 ± 0.06	46.3 ± 0.1
B	0 ±0.005	0 ±0.005	1.00 ±0.01	5.2 ±0.1	44.6 ±0.1
C	-0.362 ±0.005	-0.728 ±0.005	0.474 ± 0.006	2.5 ±0.1	24.1 ±1
D	0.941 ±0.01	-0.797 ±0.01	-	-	-
G	0.734 ±0.01	-0.649 ±0.01	-	-	-

Table 6.1: Image positions and fluxes for images A B and C measured with OSIRIS. Fluxes are in units of the [OIII] flux of image B which was measured to be $1.4 \pm 0.2 \times 10^{-14}$erg/(s cm^2) by Murayama et al. (1999). Distances are in units of arcseconds. The position for image D comes from averaging the offsets from radio observations by Patnaik et al. (1999) (see Equation 6.1), while the position for the main lens galaxy (G) comes from averaging the offsets from HST observations from the CASTLES website.

Our data give us three independent measurements of the quasar spectrum, which we expect to be magnified relative to each other due to gravitational lensing. We model the broad H-β emission and narrow emission lines as being composed as linear sums of 5th and 3rd order Gauss Hermite polynomials respectively, while the continuum is modelled as a power law. We allow more flexibility in the fit to the broad emission line to account for winds which may cause significant asymmetries in the line profile. The model imposes the same narrow line FWHM and centroid for all 3 lensed QSO images. The broad and narrow emission lines are allowed to be redshifted relative to each other to account for winds.

The amplitudes of the continuum, narrow, and broad emission lines are all free to vary between the lensed images. The broad line width and shape (i.e. the width and amplitudes of the component Gauss-Hermite polynomials which are summed to model the broad line emission), and the continuum slope are also inferred separately for each spectrum to account for the differential effects of microlensing and intrinsic variability, which will more strongly affect the continuum and high-velocity tails of the broad emission

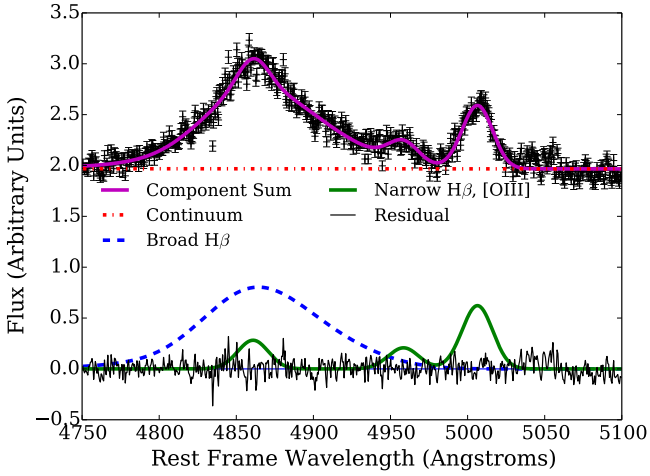

Figure 6.2: Demonstration of best fitting model spectrum to the spectrum of image B, as well as each of the separate line components.

line as these are emitted from the smallest physical area. Figure 6.2 shows an example of the decomposition of the spectrum into the four components for image B. Uncertainties in the decomposition are incorporated into the final calculated line fluxes by drawing 1000 random samples from the posterior probability distribution of the spectral parameters and re-computing the line fluxes for each sample draw. The integrated line fluxes with measurement uncertainties for each image are given in Table 6.1. In the case of [OIII], the fluxes of both lines are summed together.

6.3.1 Astrometry of image D and the main lens galaxy

Although our measurement with OSIRIS only contains information for image fluxes and positions for three of the four images of the system, we incorporate information on the position of the fourth image from previous radio imaging by Patnaik et al. (1999). The position of the fourth image aids in the constraint of the model for the main deflector,

and is not expected to vary relative to the other images over time. We take the position of image D to be the mean position after applying the offsets measured by Patnaik et al. (1999) between the three measured image positions:

$$\boldsymbol{x}_D = 1/3 \sum_{i=A,B,C} \boldsymbol{x}_i - (\boldsymbol{x}_{i,p} - \boldsymbol{x}_{D,p}), \tag{6.1}$$

where $\boldsymbol{x}_{p,i}$ represents position measurements of the i^{th} image by Patnaik et al. (1999), and $\boldsymbol{x}_i$ is the corresponding measurement in this work. The uncertainty in position D is the standard deviation of the positions. The flux of image D is not used as a constraint in the lens modelling.

Similarly, we use the relative position of the lens galaxy in HST imaging from the CfA-Arizona Space Telescope LEns Survey (CASTLES) website[2] in order to constrain the centroid of the main deflector in the lens model.

6.4 Gravitational Lens Modelling

Having obtained the integrated narrow-line fluxes and image positions, we can now test for the presence of substructure in the system using gravitational lens modelling. As noted in the introduction, the positions and fluxes of the lensed narrow-line emission images are determined by the first and second derivatives respectively of the gravitational potential of the deflector, and therefore can be used to recover the mass distribution of the perturber as well as the main deflector.

We begin in Subsection 6.4.1 by modelling the lens as containing a single main de-

[2] `www.cfa.harvard.edu/castles/`

flector with no substructure in an external shear field, and find the best fit mass model parameters given our image fluxes and positions using `gravlens` by Keeton (2001b,a). As with previous studies, we find that this mass model provides a poor fit to the image fluxes (Mao & Schneider 1998; Keeton 2001c; Bradač et al. 2002; Sluse et al. 2012a; Dobler & Keeton 2006). In Subsection 6.4.2, we explore the effects of adding a single perturber to the system with varying mass profiles.

6.4.1 Singular Isothermal Ellipsoid Deflector

Following numerous previous studies, we model the deflector as a singular isothermal ellipsoid (SIE) which has been shown to provide a good fit to the macroscopic mass distribution of lens galaxies (Treu 2010). The SIE has a density distribution given by:

$$\rho(r) = \frac{\rho_o}{r^2} = \frac{\sigma_v^2}{2\pi r^2 G} = \frac{b}{2\pi r^2}\left(\frac{c^2}{4\pi G}\frac{D_S}{D_{LS}D_L}\right) \tag{6.2}$$

Where σ_v is a velocity dispersion, b is the Einstein radius of the lens, and D_S, D_{LS} and D_L are the angular diameter distances from the observer to the source, from the lens to the source and from the observer to the lens respectively. In `gravlens`, the radius is in elliptical coordinates such that:

$$r = \sqrt{x^2(1-\epsilon) + y^2(1+\epsilon)}, \tag{6.3}$$

where ϵ is related to the axis ratio q by $q^2 = (1-\epsilon)/(1+\epsilon)$. This means that the Einstein radius is defined along the intermediate axis between the major and minor axes.

As is customary, we also allow for the presence of external shear γ in the direction θ_γ in order to account for the fact that B1422+231 is a member of a group of galaxies (e.g. Kundic et al. 1997).

We determine the lens model parameters by first using the `gravlens` optimisation routine in order to find the maximum likelihood parameters for the central deflector given the observed image positions and narrow-line flux ratios. We then use `emcee hammer` (Foreman-Mackey et al. 2013), which is a Markov Chain Monte Carlo algorithm which is efficient at exploring highly degenerate parameter spaces, in order to determine uncertainties in the parameter values. We initialise the proposal values in a small region in parameter space around the best fit solutions from `gravlens`, and then for each proposed set of host properties, we use `gravlens` in order to optimise the source position and compute the χ^2 for the resulting image positions and magnifications.

The inferred median and one sigma uncertainties for the lens model parameters are given in Table 6.2. The corresponding model prediction for the image fluxes and positions are given in Table 6.3. The model provides an excellent fit to the image positions, matching them within the measurement uncertainties, however the predicted image flux ratios deviate significantly from the observation, so that the typical model χ^2 is ~ 50 for seven degrees of freedom. As an additional test, we repeat the inference, this time not including image fluxes as a constraint (row two of Table 6.3). In this case, the fit is much better, with a χ^2 of $\sim$1 for four degrees of freedom. The inferred model parameters in this case are consistent with previous studies which used IR and radio data but did not include flux information in their lens models (Mao & Schneider 1998; Keeton 2001c; Bradač et al. 2002; Sluse et al. 2012a).

The significant deviation between the observed [OIII] image flux and the smooth model prediction is consistent with previous observations of the system across a broad range of wavelengths, which also found that flux of image A is about 20% brighter

relative to image B than predicted by the single deflector model (Mao & Schneider 1998; Keeton 2001c; Bradač et al. 2002; Sluse et al. 2012a; Dobler & Keeton 2006). Figure 6.3 compares the [OIII] flux ratios between images A, B and C with measurements made at other wavelengths, and also relative to the smooth model prediction.

In Figure 6.3, we also show our measured flux ratios in the continuum and broad Hβ emission. The broad Hβ emission has flux ratios consistent with that of the narrow emission, while the continuum emission deviates more dramatically from the smooth model prediction. This is consistent with microlensing which can act in conjunction with millilensing by a subhalo, and would most strongly affect the smaller continuum source size. Note that while the macromodel can be adjusted to provide a reasonable fit to the C/B flux ratios at other wavelengths, this is not possible in the case of the A/B flux ratios.

In order to improve the model fit, a natural solution is to add a perturber less massive than the central galaxy. This can alter image fluxes without significantly shifting image positions, because image fluxes depend on the second derivative of the lensing potential, while image positions depend on the first derivative. In the following subsection we explore the effects of adding a single perturbing mass to the lens system.

6.4.2 Smooth model plus Perturber

As can be seen in Figure 6.3, the ratio between the fluxes of images A and B deviates significantly from the smooth model prediction while the ratio between fluxes of image C and B does not. The simplest solution is that image A has been magnified by a nearby

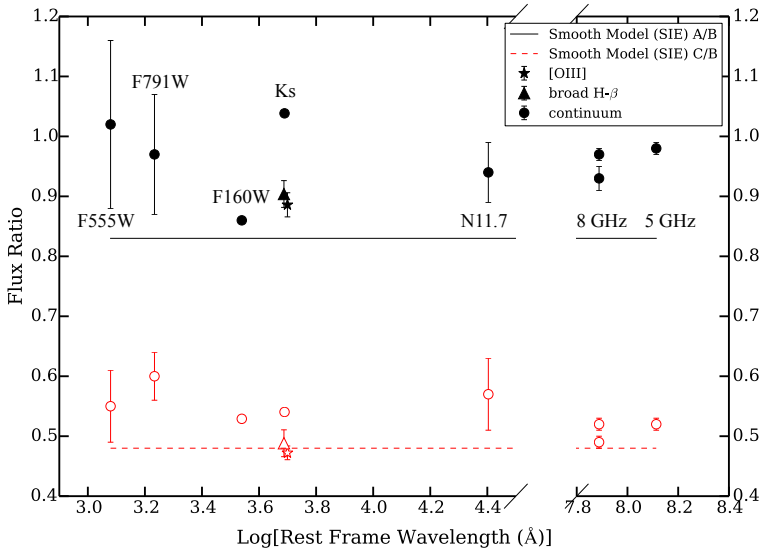

Figure 6.3: Comparison of measured ratios of images A B and C as a function of rest frame wavelength, with the narrow (star), broad (triangles) nd continuum Ks measurement from this work. Continuum measurements are labelled with the observed filter for convenience. We include HST measurements with WFPC2 F791W and F555W measurements from the CASTLES website, while the NICMOS F160Wa measurement is from Sluse et al. (2012a). The mid-IR N11.7 measurement is from Chiba et al. (2005). Radio measurements are from Patnaik et al. (1992, 1999). The solid and dashed lines represent the best-fit prediction to the [OIII] fluxes for the flux ratios between images A/B and C/B respectively, assuming one lens galaxy which is a singular isothermal ellipsoid with external shear. While the ratio between image B and C agrees well with the model, the ratio between A/B is significantly different from the smooth model prediction.

perturber (Mao & Schneider 1998; Dobler & Keeton 2006; Bradač et al. 2002), with the amount of magnification depending on the perturber position, mass scale and mass profile.

We explore the simple case of adding a single perturber to the SIE plus external shear model. This is a useful way to understand how our astrometric and photometric precision allows us to constrain the presence of substructure, and to compare our results with previous works, which have also assumed a single perturbing subhalo to the smooth lens model. In reality, the system is likely to contain thousands of lower mass subhalos. However the majority of these subhalos will be too low in mass and too far from the lensed images to affect them. The degeneracies between perturber positions and mass make it difficult to distinguish between the effects of populations of perturbers and a single perturber added to the main lens system (Fadely & Keeton 2012). With a larger sample of lens systems, the statistical properties of the perturber population can be inferred more robustly, thus we will leave a more physical model with a realistic population of substructure to a future work in which we consider a larger sample of lensed quasar systems.

For illustrative purposes we consider three mass profiles for the perturber, in order to explore how the mass profile affects the perturber lensing signal. Below we discuss each of the mass profiles in detail.

The simplest mass profile computationally is the singular isothermal sphere (SIS) given by Equation 6.2, with ellipticity set to zero. This profile is not physical as it diverges as r goes to zero and extends outward indefinitely, however it provides a good approximation to the small scale density profiles of galaxies (e.g. Gavazzi et al. 2007;

Koopmans et al. 2009a; Lagattuta et al. 2010).

It is expected that a subhalo will undergo truncation by tidal forces from the main halo. This truncation can be modelled by a pseudo-Jaffe profile (PJ, Muñoz et al. 2001), which has a density distribution:

$$\rho(r) = \frac{\rho_o}{r^2}\frac{a^4}{r^2 + a^2}, \tag{6.4}$$

where ρ_o is the same as in Equation 6.2. The truncation radius, a, is traditionally calculated by assuming that the perturber is exactly within the plane of the lens galaxy at the Einstein radius, b_{host}, so that $a = \sqrt{b_{\rm sub} b_{\rm host}}$ (Metcalf & Madau 2001). In reality, tidal stripping depends on a variety of factors, including the orbit of the perturber around the host galaxy. This assumption for the tidal truncation radius assumes that the current projected position is also the pericentre of the satellite orbit. The SIS and PJ models bracket two extreme cases of no tidal stripping on the one hand, and maximum tidal stripping (in the absence of baryons) on the other.

The third mass profile we consider is the Navarro Frenk and White (Navarro et al. 1996, NFW) profile. This mass profile is predicted by dark matter-only simulations, which may be a better match to the mass profile of subhalos than the SIS profiles. In fact the mass profile of subhalos may be even shallower than this at the centre (e.g. Walker & Peñarrubia 2011; Wolf & Bullock 2012; Hayashi & Chiba 2012), but we leave a more complete exploration of the effects of mass profile variations to a future work. The density distribution of the NFW halo is:

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2} \tag{6.5}$$

Where r_s and ρ_s are the scale radius and the density at the scale radius respectively. To

reduce the number of free parameters, we apply the mass-concentration relation predicted by Macciò et al. (2008), assuming a WMAP5 cosmology (Dunkley et al. 2009). This is a large extrapolation at dwarf galaxy scales, but allows us to make a useful one-parameter comparison with the steeper mass profiles.

As a first step, we again use `gravlens` to find the best-fit model to the narrow-line fluxes and positions for the macromodel parameters (host and external shear) as well as the perturber mass and position in the case of each of the three perturber mass profiles. As with the SIE only model, we then use `emcee hammer` to infer the model parameters, allowing both the macromodel and the substructure parameters to vary, assuming a uniform prior on the perturber position and log-uniform prior on the perturber mass. For each proposed perturber mass scale, position and set of host properties, we use `gravlens` in order to optimise the source position and compute the χ^2 for the resulting image positions and magnifications.

As we will demonstrate below, despite the degeneracies, the possible perturber positions are restricted by the proximity of image B to image A, as well as by the critical curve of the main lens. Furthermore, the astrometric precision of the measurement prevents a subhalo large enough to cause significant astrometric deviations from the smooth model. These effects can all be accounted for in the `gravlens` code. We also place lower limits on the mass of the perturber based on the finite size of the source, which we describe below.

6.4.3 Finite Source Effects

The finite nature of the source places lower limit to the perturber mass that can cause the observed magnification (Dobke & King 2006). This can qualitatively be understood by the fact that observed magnification is the convolution of the magnification pattern with the source surface brightness distribution. The larger the source the more small-scale features in the magnification due to low-mass substructure will be smeared out. In the following subsection we discuss how we infer the narrow-line emission source size and use this to place lower limits on the perturber mass for the three different perturber mass profiles.

The observed images in each exposure are a convolution of the PSF in that exposure which we modelled as the sum of two concentric Gaussians in Section 6.2, with the source which we assumed to be a point source. This was a valid assumption given that the flux in the images between 4500-5100 Å is dominated by continuum and broad H$-\beta$ emission, which are point-like (μas) relative to the resolution of the telescope. Narrow-line emission, on the other hand, is seen to extend out to hundreds or even thousands of parsecs for the most luminous Seyferts at low redshift (e.g. Bennert et al. 2002), although the dominant contribution to the narrow-line luminosity comes from the central tens of parsecs (Müller-Sánchez et al. 2011).

In order to determine the size of the lensed narrow-line source, we adjust the analysis described in Section 6.2, this time modelling the flux as being emitted from a Gaussian light distribution. This Gaussian source, when convolved with the PSF Gaussian yields

observed image widths given by

$$\sigma_{eff} = \sqrt{\sigma_{PSF}^2 + (\mu^{1/2}\sigma_{source})^2}, \tag{6.6}$$

where σ_{PSF}^2 is the standard deviation of the seeing-halo, and μ is the best fit magnification given by `gravlens` at each of the image positions.

We apply this analysis to a narrow-line image, created by subtracting a variance weighted mean image of the continuum region between rest-frame 5070-5140 Å from a variance weighted mean image which contains [OIII] plus continuum between rest frame 4980-5030 Å. We infer an intrinsic source size of $\sim$ 15 mas. This analysis assumes that the diffraction core of the AO corrected PSF had a fixed size of 6.5 mas, while in reality the size of the diffraction core depends on the AO performance and weather conditions, and can vary from night to night. To test this, we also infer the source size in an image containing light from only the quasar continuum, which we expect to be unresolved. Instead, we find that the continuum image has an intrinsic source size of $\sim$ 10 mas, implying that the diffraction core of the PSF is not as narrow in reality as in our model. However, the [OIII] emission is definitely larger and thus resolved. Subtracting the two in quadrature we estimate the intrinsic [OIII] size to be roughly $\sim$ 10 mas. This corresponds to approximately 60 pc at the redshift of the quasar.

Based on the size-luminosity relation for type-I AGN measured in the Local Universe by Bennert et al. (2002), we would have expected the full narrow-line region in B1422+231 to extend to several hundreds of pc. However, we are likely to only be sensitive to the highest surface brightness parts of the narrow line regions because of surface brightness sensitivity and cosmological dimming. So it is expected that our measure-

ment be smaller than in the local universe. Furthermore, we do not consider lensing shear distortions which are expected to cause the images to be elongated tangentially to the radial direction, rather than simply enlarged isotropically as our analysis assumes. Thus this simple calculation should be considered a lower limit to the true narrow line emission region size, which corresponds to a conservative lower limit to the perturber mass. In other words, a larger narrow line region would result in a more narrow posterior distribution function of the perturber mass.

Dobler & Keeton (2006), computed the relationship between the source size and the minimum Einstein radius of an SIS perturber that can cause the observed magnification of image A, given the convergence and shear of the system B1422+231. They found that an SIS perturber must have mass within the Einstein of $b > 0.056a$, where b is the Einstein radius of the SIS perturber, and a is the unlensed source size. This corresponds to a lower mass limit of $\sim 10^{6.5}$ and $\sim 10^{6}$ $M_{\odot}$ within 600 pc in the case of the SIS and PJ perturbers respectively, assuming the perturbers are in the plane of the lens galaxy[3]

The mass limit on the NFW perturber cannot be computed analytically. We approximate this limit by computing the average magnification in a grid of points with a size of the magnified image, for increasing perturber masses. The limit is then given approximately by the minimum perturber mass that can achieve the observed magnification in a region the size of the magnified image. This is an approximate result, as it assumes the source is magnified isotropically rather than distorted as it actually is. In the case of the SIS perturber, for instance, with this method we find that the perturber must

[3]The interpretation of the enclosed mass given a fixed Einstein radius and density profile depends on the perturber redshift.

have $b > 0.04a$, which is slightly lower than the true value computed by Dobler & Keeton (2006), but is good enough for an order of magnitude estimate of the minimum perturber mass. In the case of an NFW perturber with mass concentration relation given by Macciò et al. (2008), the perturber must have a scale radius which is approximately five times larger than the unlensed source size, which corresponds to a mass limit of $\sim 10^{6.5} M_{\odot}$ within 600 pc.

6.5 Results

The SIE host plus perturber lens models show significant improvement relative to the SIE-only model for all three perturber mass profiles, with best fit χ^2 of one or smaller for four degrees of freedom. The model prediction for the image positions and fluxes are listed in Table 6.3, for the case of an SIE host with an SIS perturber. The model parameters which determine the properties of the main SIE lens and external shear field are consistent for the SIE plus perturber model and the SIE only model when fluxes were not included as a model constraint. This illustrates how the perturber can alter image fluxes without significantly affecting image positions.

Model	host b	ϵ	host PA	γ	θ_γ	$\log_{10}[b_{SIS}]$	dRA_{SIS}	$dDec_{SIS}$	χ^2/DOF
SIE (fluxes)	0.746±0.007	0.36±0.03	-62±1	0.16±0.01	-47±2	-	-	-	7
SIE (no fluxes)	$0.771^{+0.007}_{-0.009}$	0.16±0.08	-57$^{\pm}$5	0.22±0.03	-54±1	-	-	-	-
SIE + SIS (fluxes)	0.765±0.009	0.17$^{\pm}$0.07	-58±5	0.22±0.02	-54^{+2}_{-1}	$-1.9^{+0.6}_{-0.7}$	$0.5^{+0.1}_{-0.2}$	0.4±0.2	0.1

Table 6.2: Posterior median and 68% confidence intervals for lens model parameters, and maximum likelihood χ^2 per degree of freedom for the two cases in which the full data set is used. In the first two cases, the lens was modelled as an SIE lens with external shear, first using image fluxes and positions as well as the galaxy position to constrain the model, in the second case using only the image and galaxy positions to constrain the model, for comparison with previous works. In the third row, the lens is modelled as an SIE in external shear with a perturbing subhalo with an SIS mass profile, using image fluxes and positions, and the galaxy position as constraints. Results are inferred based on the [OIII] image fluxes and positions. Parameters are defined in Section 6.4.1. Position angle and shear are in units of degrees East of North while Einstein radii are in units of arcseconds.

Image	dRa_1	$dDec_1$	$Flux_1$	dRa_2	$dDec_2$	$Flux_2$
A	0.375 ±0.003	0.310±0.004	0.83 ± 0.01	0.376 ± 0.004	0.327 ± 0.005	0.88 ± 0.02
B	0.002 ±0.003	0.014 ±0.004	1.00	0.000±0.004	0.000 ±0.005	1.00
C	- 0.346 ±0.005	-0.743 ±0.003	0.484 ± 0.006	-0.339 ±0.005	-0.739 ±0.004	0.471 ± 0.007

Table 6.3: Posterior median and 68% confidence interval for model predicted image positions in units of arcseconds and fluxes in the case of a single smooth lens galaxy (1) and a lens galaxy plus perturber with SIS mass profile (2). Results for the other two smooth plus perturber mass profiles are similar.

In Table 6.2 we list the median and one sigma confidence intervals for the lens model parameters in the case of an SIE host galaxy and an SIS perturber. The host model parameters are consistent with those in the smooth model case, which is expected given the fact that the perturbing halo is a relatively small addition to the mass of the system.

It is useful to quantify the significance of the perturbation to the smooth model. Here we discuss three different model testing methods. First, the inferred distribution for the perturber position and mass is informative given that the SIE plus perturber models allows for solutions in which the perturber does not contribute significantly to the lensing, for instance if it were low mass and far from the lensed images. In this sense the more complex SIE plus perturber model also includes the simpler possibility of the SIE-only model. We find that in 1.8 million iterations of the MCMC, the perturber is always located such that it is contributing significantly to the magnification of image A relative to the smooth model at that location.

In addition, the maximum likelihood χ^2 per degree of freedom for the SIE-only model is 7 and the for SIE plus SIS perturber model it is 0.1. Results are similar for the other two perturber mass profiles. An alternative way to compare models is via the Akaike Information Criterion, (Akaike 1974) modified to take into account the relatively large number of model parameters relative to data in this case (AICc Hurvich & Tsai 1989). This is useful when trying to determine whether the addition of new parameters to a model improves the fit relative to the 'true' underlying model (Kelly et al. 2014).

The AICc is calculated as:

$$AICc = 2k - 2\log p_{\rm ml} + \frac{2k(k+1)}{n-k-1} \tag{6.7}$$

Where here k is the number of model parameters, $p_{\rm ml}$ is the maximum likelihood fit to the data given a model choice, and n is the number of independent data points used to constrain the model. The AICc is similar to the χ^2 with an additional penalty for extra model parameters. Given maximum $\log p$ values of the SIE only and SIE plus SIS models, the AICc values are 60 and 50. This indicates that the SIE plus SIS model is of order 100 times more likely than the SIE only model.

While all three perturber mass profiles provide equally good fits to the observations, the posterior probability distributions for the perturber's position and mass vary significantly as can be seen in Figure 6.4. There is similar qualitative behaviour in the sense that the further a perturber is from image A, the more massive it must be to achieve the same lensing effect. Furthermore, in all three cases, the mass range is limited naturally by the astrometric precision of the measurement. If the perturber becomes too massive it causes significant astrometric perturbations to the whole system, rather than just affecting the flux of image A. However, the posterior probability distribution for the SIS perturber position is much broader than in the case of the PJ profile. This is due to the fact that the SIS mass profile extends outward indefinitely, so that it has a fundamentally different lensing effect than the PJ perturber outside of the truncation radius of the PJ perturber. We discuss this in more detail in the Appendix. Another interesting feature is that the shallower NFW profile is not restricted by the position of the lens critical curves in the way the SIS and PJ perturbers are. This shallower profile integrates to a larger total mass for fixed aperture mass. Given that the lensing effect is determined by the aperture rather than total mass, this implies that to achieve the same magnification as the SIS or PJ perturbers, the total mass of the NFW perturber must be higher. This

limits the 95% and 68% position contours for the NFW profile to be closer to image A than the other two profiles.

In Figure 6.5, we plot the marginalised posterior probability distribution for the perturber mass within 600 parsecs, assuming that the perturber is located in the plane of the lens galaxy. We note that for PJ perturbers with masses lower than $\sim 10^{8.5}$, the truncation radius falls inside of 600 pc. The 68% confidence intervals for the logarithm of the perturber mass within 600 pc are $8.2^{+0.6}_{-0.8}$, $8.2^{+0.6}_{-1}$ and $7.6^{+0.3}_{-0.3}$ $\log_{10}[M_{sub}/M_{\odot}]$, for a singular isothermal sphere, PJ, and NFW mass profile respectively. The PJ perturber can have lower masses relative to the SIS mass profile due to the truncation radius which simulates tidal stripping, and falls within 600 pc for PJ perturbers with masses less than $\sim 10^{8.5} M_{\odot}$ within 600 pc. The NFW profile is restricted to a small range of masses due to its shallow density profile which gives it an effectively weaker lensing signal. For ease of comparison with Local Group studies, the mass within 300 parsecs is shown in the upper panel of Figure 6.5, and has a 68% confidence interval of $7.8^{+0.6}_{-0.7}$,$8.0^{+0.6}_{-0.8}$, $7.2^{+0.2}_{-0.2}$ $\log_{10}[M_{sub}/M_{\odot}]$ for the three mass profiles.

Given the significantly different lensing effects in the three cases, it will be possible with a larger sample of lenses to learn about both the typical mass profile as well as the mass function of the perturbers.

6.6 Discussion

Our primary goal in this work is to demonstrate that strongly lensed narrow-line emission provides an alternative to radio emission in lensed quasars for detecting substructure,

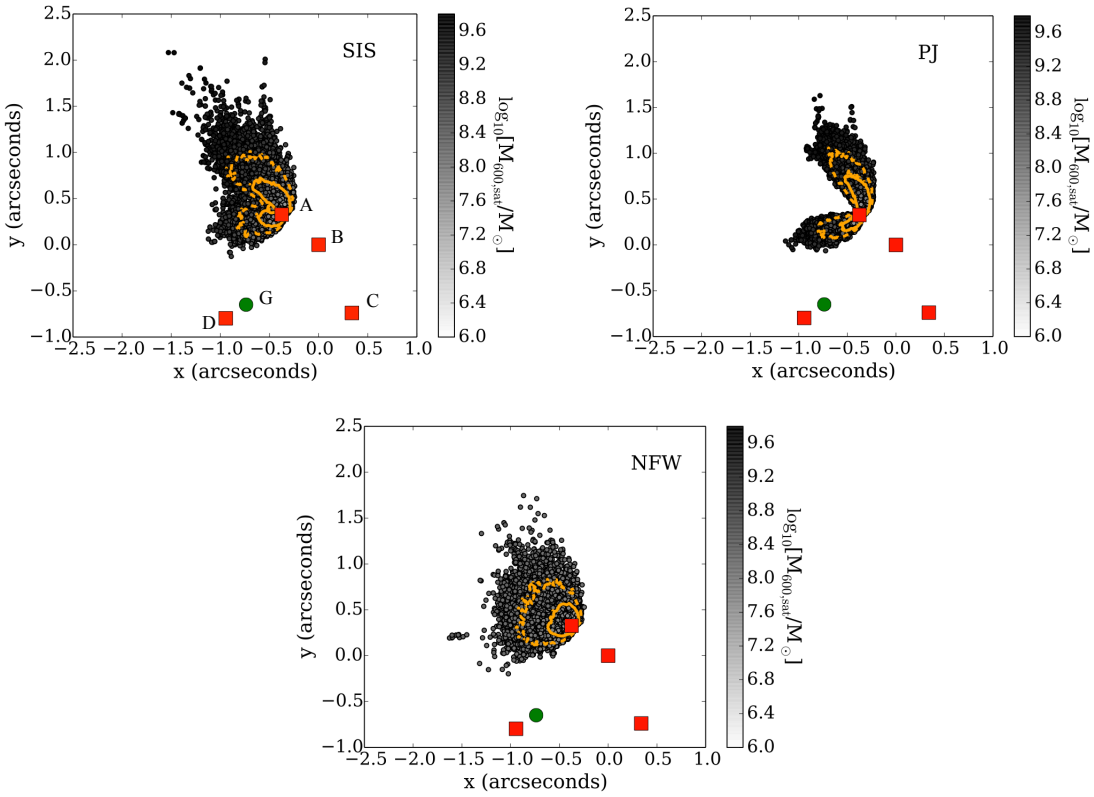

Figure 6.4: Posterior probability distributions of the perturber position relative to the lensed images shown as red squares, and lens galaxy shown as a green circle, for a single SIS, PJ, and NFW perturber from top to bottom. The grey scale represents the perturber mass within 600 pc assuming the perturber is in the plane of the lens galaxy, and solid and dashed contours represent the 68 and 95% confidence contours respectively relative to the most likely position.

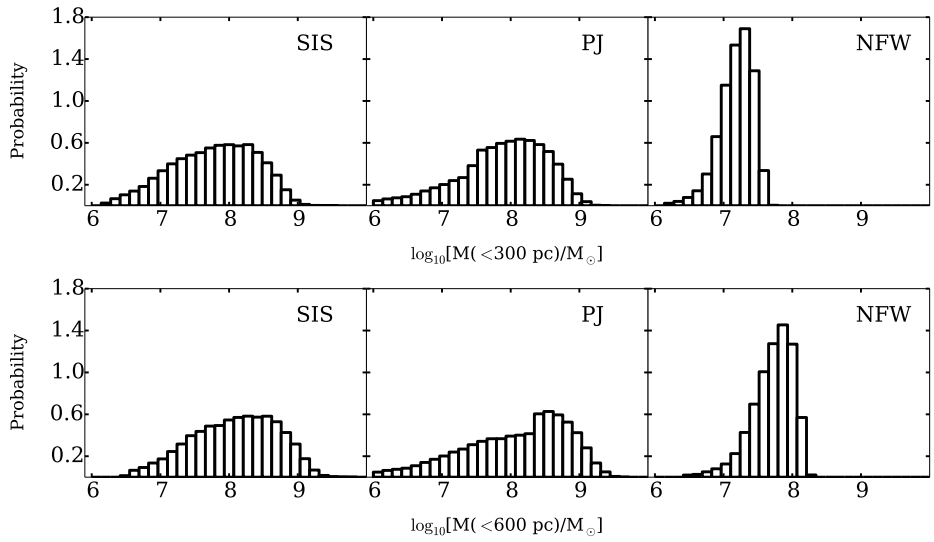

Figure 6.5: Marginalised posterior probability distribution of the perturber mass within 300 pc (upper row) and 600 pc (lower row) assuming an SIS, PJ and NFW mass profile respectively from left to right.

yielding lens model constraints with consistent results and comparable accuracy. In this section we discuss previous results for this system, how the analysis can be improved by considering additional systematic uncertainties, as well as future prospects for the method

6.6.1 Comparison with previous work

B1422+231 is a bright, radio loud system which enabled one of the earliest detections of dark substructure at cosmological distances (e.g. Mao & Schneider 1998; Keeton 2001c; Bradač et al. 2002). These studies found that the observed radio fluxes could be best explained by a perturbing substructure with characteristic Einstein radius of a few mas. Although there was some question about how macromodel assumptions might affect the inference of substructure (Kawano et al. 2004), Kochanek & Dalal (2004) demonstrated

that smooth macromodels that are consistent with observations of galaxies using weak lensing and galaxy morphology do not provide a good fit to the observed image positions and flux ratios. Our analysis based on lensed narrow-line emission yields consistent results to these studies.

6.6.2 Limitations of our analysis

We have explored how our data can be used to constrain a mass model with a main lens in external shear and one perturbing subhalo. As we discussed in Section 6.4.2, this model makes several important simplifying assumptions.

First, for simplicity we adopt a macro-model composed by a singular isothermal ellipsoid with external shear. This is in general found to be a good model for early-type galaxies both in terms of radial profile Koopmans et al. (2006, 2009b) and angular structure (Yoo et al. 2005, 2006). However, allowing for more flexibility in the model could inflate the uncertainties on the inferred substructure, or reduce its significance, even though in the case of B1422+231 it is hard to find physically plausible models that fit the data without a localized substructure (Evans & Witt 2003).

Second, as with previous works, we explore the effects of a single perturbing subhalo (MacLeod et al. 2009; McKean et al. 2007; MacLeod et al. 2013). When studying the effects of subhalos on lensed compact image fluxes and positions, full populations of subhalos are computationally expensive and not well constrained by a single lensing system (Fadely & Keeton 2012). In the future, when large samples will be available, it will be important to carry out a systematic analysis of the entire population of subhalos

using hierarchical modelling to infer at once the mass function and spatial distribution of the subhalos.

Third, similarly to other works of this type, we have assumed that the perturbing substructure is in the plane of lens galaxy in order to estimate the perturber mass (e.g. Dalal & Kochanek 2002; MacLeod et al. 2009; Fadely & Keeton 2012; Vegetti et al. 2010b, 2012; Vegetti et al. 2014). Xu et al. (2012) showed that line of sight structure may contribute significantly to image magnifications, although the observed frequency of radio flux ratio anomalies can be explained with in situ perturbers predicted by CDM (Xu et al. 2013b; Metcalf & Amara 2012). This has important implications for the inferred subhalo mass function as all masses are inferred assuming the redshift of the perturber is known. The lensing effect varies with mass if the perturber is not within the plane of the lens galaxy (Xu et al. 2012; McCully et al. 2014b).

We have also demonstrated the effects of varying the assumed mass profile of the perturbing subhalo, which significantly change the posterior probability distribution for the perturber mass and position. It would be interesting in a future work to consider a range of subhalo mass profiles which are drawn directly from simulations, so that differential tidal stripping as a function of three dimensional distance from the lens halo centre, and scatter in the subhalo concentration mass relation can be taken into account for instance.

6.6.3 Future Prospects

Many more quasars have significant narrow-line emission than radio emission, thus this method can be extended to study substructure in a larger sample of systems. There is significant future potential for this method as ongoing and planned optical surveys such as DES, LSST, PANSTARRS and GAIA are expected to find thousands more lensed quasars based on their optical rather than radio properties (Oguri & Marshall 2010). The next generation of adaptive optics systems (NGAO), (Max et al. 2008), and telescopes such as the Thirty Meter Telescope and the European Extremely Large Telescope will make it possible to measure narrow-line lensing rapidly and with high precision for those large samples of objects. In the nearer future, upgrades to the Keck AO system will enable PSF telemetry which will thereby make it possible to obtain better constraints on the source size, and thus place stronger lower limits on perturber masses.

Significant information is also contained in deep imaging of the lensed systems, which can sometimes reveal luminous substructure (McKean et al. 2007; MacLeod et al. 2009). If a luminous satellite is detected, it can be used to break the degeneracy between the subhalo mass and position, as well as providing a measurement of the subhalo mass to light ratio.

6.7 Summary

We used OSIRIS at Keck with adaptive optics in order to obtain spatially resolved spectroscopy of the gravitational lens B1422+231. We developed a new pipeline which enabled us to robustly extract lensed image spectra, by first inferring the PSF and image

positions in each exposure. Using this information, we detected a significant deviation from a smooth gravitational lens model, which we used to infer the presence of a perturbing low mass subhalo. Our main results are summarised below:

1. Using our data reduction pipeline, we were able to measure image positions to 5 mas accuracy and integrated line flux ratios with $\sim$3% uncertainties.

2. The narrow-line flux ratios are consistent with radio measurements, and deviate significantly from the smooth model predictions. The broad-line flux ratios are consistent with the narrow line measurements. Our measurement of the continuum is significantly offset, showing a much stronger deviation from the smooth model, possibly indicative of microlensing in addition to millilensing.

3. Based on the assumption that the deviation of observed narrow-line fluxes from the smooth model is due to a single perturbing subhalo, we infer the mass and position of the perturbing subhalo for three different mass profiles. The mass profile significantly affects possible perturber positions and masses, with an NFW subhalo restricted to the smallest region of parameter space by its shallower mass profile which has a relatively weaker lensing signal. The 68% confidence intervals for the logarithmic perturber mass within 600 pc are: $8.2^{+0.6}_{-0.8}$, $8.2^{+0.6}_{-1}$ and 7.6 ± 0.3 $\log_{10}[M_{sub}/M_{\odot}]$, for a singular isothermal sphere, pseudo-Jaffe, and NFW mass profile respectively. The mass within 300 pc is $7.8^{+0.6}_{-0.7}$ and $8.0^{+0.6}_{-0.8}$, $7.2^{+0.2}_{-0.2}\log_{10}[M_{sub}/M_{\odot}$.

Acknowledgements We thank B. Kelly, M. Bradac, S. Vegetti, D. Sluse, C. Keeton, S. Mao, G. Dobler, P. J. Marshall, B. Brewer, V. Bennert, A. Pancoast, S. Suyu

and C. Kochanek for useful comments and discussion. We also thank support astronomer S. Dahm and observing assistant T. Stickel. AMN and TT acknowledge support by the NSF through CAREER award NSF-0642621, and by the Packard Foundation through a Packard Fellowship. AMN acknowledges support from a UCSB Dean's Fellowship.

6.8 Appendix: Variation in lensing effect with perturber mass profile

In this appendix we illustrate the difference in the posterior probability distributions for the perturber position and mass seen in Figure 6.4 in the case of the PJ and SIS perturber, by means of a simple example. The most notable difference between the two models is the large gap to the East of image A in the case of the PJ perturber. In order to clarify the origin of this difference, in Figure 6.6 we show the model prediction for the image positions as the perturber Einstein radius is varied for fixed position and host macromodel, in the case that the perturber is in the gap (position 2), and also in the region further North which is permitted for both mass profiles (position 1). We also show the results in the case of a point mass perturber, which has no convergence at all at the image positions. This is the extreme opposite to the SIS, which has a density profile which extends outward indefinitely. The perturber positions relative to the lensed images are given in the lower left panel of Figure 6.6.

The first thing to note in this figure is that the perturber in position 1 does not provide a good fit to the observation in either the case of the SIS or PJ perturber (the

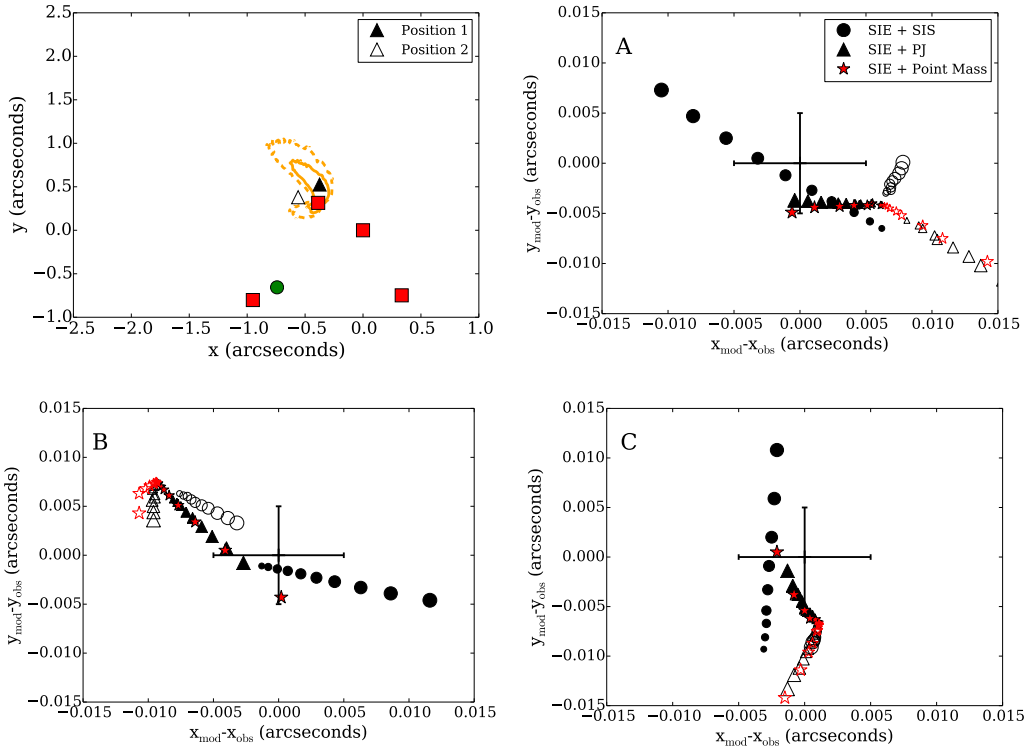

Figure 6.6: Demonstration of how the model image positions vary relative to the observed image positions as the perturber Einstein radius is varied logarithmically for fixed perturber position and SIE macromodel parameters, for an SIS, PJ and point mass perturber represented by circles, triangles, and red stars respectively. Point sizes increase with increasing perturber Einstein radius. We explore two perturber positions, illustrated in the upper left panel by open and filled triangles, where the red squares and green circle represent observed image and galaxy positions respectively and the solid and dashed contours represent the 68% and 95% confidence interval for the position of the perturber in the case of the PJ mass profile.

best fit χ^2 per degree of freedom is approximately 8 in the case of the SIS perturber, and ~ 50 in the case of the PJ perturber). This figure illustrates the fundamental difference in the lensing effect between the truncated and non-truncated perturber profiles, due to the fact that the lensing signal in the case of the truncated profile is dominated by shear, while the SIS profile has significant convergence at the image position. In fact, the PJ perturber has an almost identical effect to the point mass perturber. In position 2, the PJ profile simply cannot provide a good fit to the astrometry, irrespective of its mass.

In Figure 6.7 we plot the resulting flux ratios as the perturber masses are varied in the three cases. As before, we see that in the three cases, the perturber in position 2 provides a less good fit to the observation in the case of the SIS. The fit is much worse though in the case of the PJ and point mass perturbers which again show almost identical behaviour. Unlike the case of the SIS profile, in position 2 the flux of image A relative to image B does not increase significantly as the perturber mass is increased for the PJ perturber. As expected, the flux ratio between image B and C does not vary significantly as the perturber mass is varied, since the dominant effect is to alter the flux of image A.

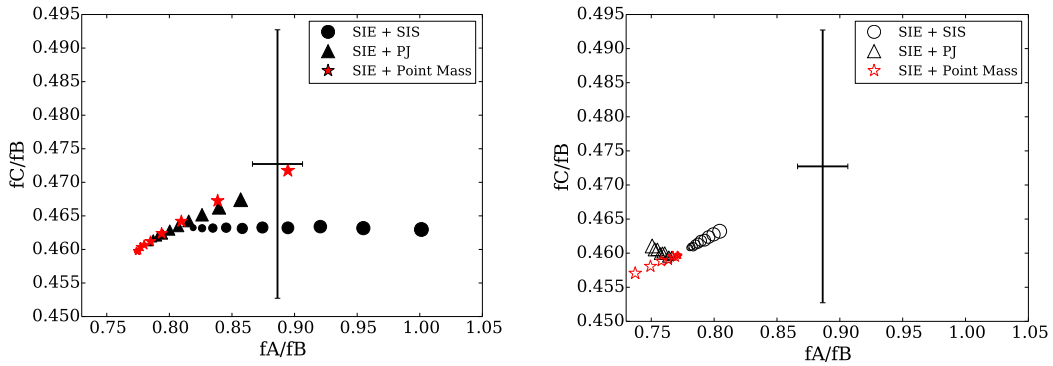

Figure 6.7: Model flux ratios as the perturber mass is varied for fixed position for two different mass profiles in the same way as in Figure 6.6. For clarity, we show results for positions one and two separately.

Chapter 7

Conclusions and future directions

This thesis explores how observations of satellite galaxies can be used to constrain ΛCDM at sub-galaxy scales, and to test theories of galaxy formation in low mass halos. There is significant future potential for the methods discussed here. In particular, we have highlighted the importance of studying satellite galaxies outside of the Local Group, as significant physical insight can be gained by studying the evolution of satellite galaxies, and how their luminosity function varies with the properties of the host galaxy. In Chapter 4, we discussed how significant future insight can be gained by studying the *colors* of satellite galaxies at a range of redshifts and mass scales. In Chapter 6, we discussed how narrow-line lensing could be applied to a much larger sample of systems, thereby enabling a more sophisticated gravitational lensing constraint on the subhalo mass function.

In this Chapter, we discuss future plans in more detail. In Section 7.1 we discuss how multi-band imaging from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS, Grogin et al. 2011; Koekemoer et al. 2011) WFC3 survey will enable a measurement of the colors and stellar mass function of satellite galaxies, as well as a deeper measurement of the satellite luminosity function. In Section 7.2, we outline future

prospects for the narrow-line lensing program. Finally in 7.3 we discuss how results from the two research programs can be combined.

7.1 The colors and stellar mass function of satellite galaxies

The CANDELS survey is a large Wide Field Camera 3 (WFC3) survey which covers five legacy fields which have extensive prior multi-band HST and ground based imaging and spectroscopy. Most exciting for the study of satellite galaxies is the near-IR CANDELS imaging with F125W, F140W and F160W which covers ~ 945 square arcminutes. This deep near-IR imaging is very sensitive, enabling the detection of fainter satellite galaxies than was possible with the F850LP or F814W in Chapters 2 and 3. Figure 7.1 shows the limiting stellar mass sensitivity of the CANDELS fields, estimated using a range of possible galaxy spectral energy distributions, as well as the predicted number of satellites and host galaxies that we will detect based on abundance matching predictions from (Busha et al. 2011). Combining the wide and deep fields, we expect this survey to contain approximately 2000 host-satellite systems with host stellar mass greater than $10^{10.5}M_{\odot}$ (the minimum host mass criterion in Chapter 3) between redshifts 0.1 and 1.4. In this section we discuss our ongoing measurement of the colors and luminosity function of satellite galaxies in CANDELS, as well as future work we plan to undertake, measuring the satellite stellar mass function.

Figure 7.2 shows the satellite detection limits we expect in the CANDELS fields

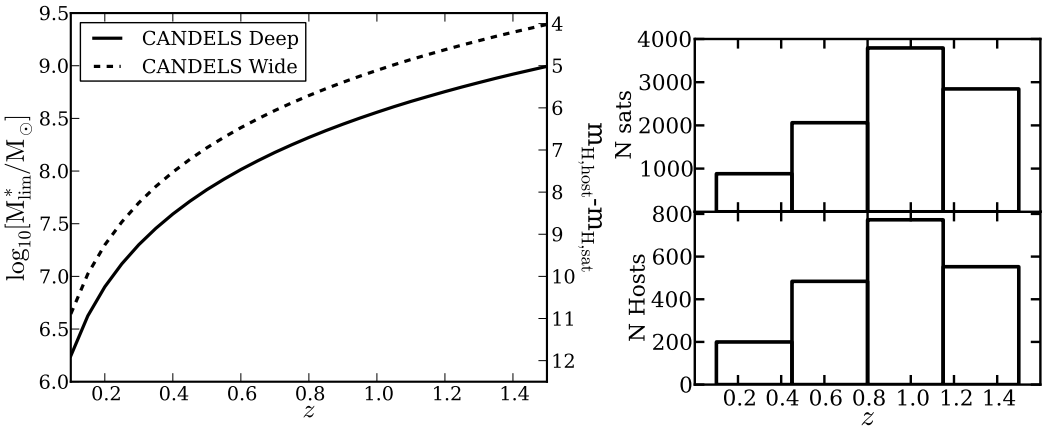

Figure 7.1: **Left**: The limiting stellar mass sensitivity of the CANDELS wide and deep fields, estimated using a range of galaxy spectral energy distributions. The y axis on the right shows the magnitude difference between the lowest mass detectable satellite and a host galaxy with stellar mass of $10^{11} M_\odot$. **Right**: Predicted number of hosts and satellites with CANDELS imaging as a function of redshift. Satellite numbers are estimated using subhalo abundance matching predictions from (Busha et al. 2011)

based on the stellar mass sensitivity of the survey, relative to the COSMOS observations from Chapter 3 and the theoretical predictions discussed in Chapter 4. By measuring the luminosity function of satellites in CANDELS, we will gain significant additional constraining power between these and other theoretical models for the satellite luminosity function, because these models vary the most for fainter satellites and at higher redshifts.

In addition to a much deeper and higher redshift measurement of the satellite luminosity function, the CANDELS survey will enable the measurement of satellite colors, thereby making it possible to directly test the theoretical predictions for the satellite color distribution discussed at the end of Chapter 4. We will use the multi-band imaging in order to measure satellite colors. In Figure 7.3, we show the color predictions from Figure 4.5, now in HST filters chosen to straddle 4000 Åbreak. Clearly the models predict very different color distributions for the satellite galaxies as they did in the rest frame u-i

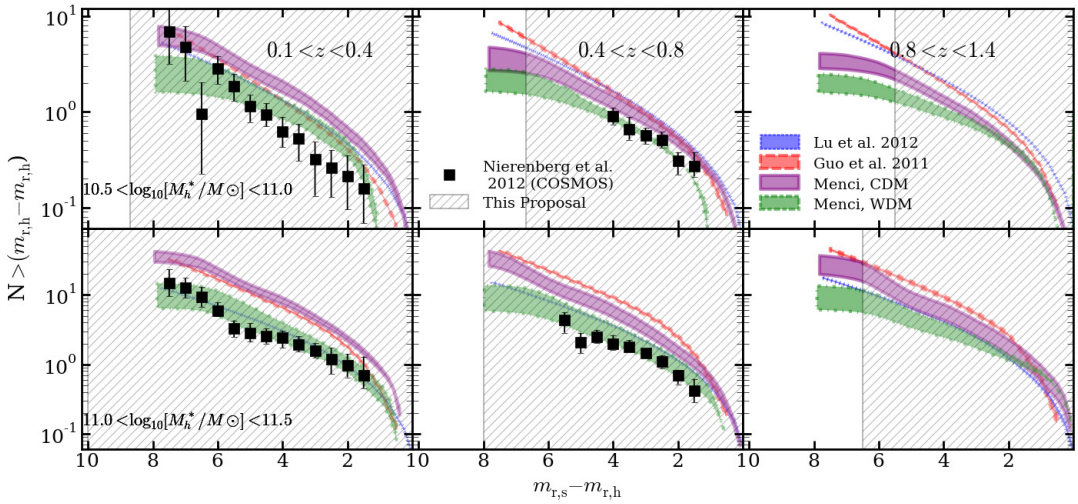

Figure 7.2: Comparison of the number of predicted and observed cumulative number of satellites as a function of the difference between host SDSS−r magnitude ($m_{r,h}$) and satellite magnitude ($m_{r,s}$), over a range of redshift and host stellar masses. Black points with error bars, and the purple blue and red curves are the same as in Figure 4.1. The hatched regions indicate the satellite luminosity range accessible in the CANDELS fields. This data makes it possible to test these and other star formation models at unprecedented magnitudes and redshifts.

colors in Figure 4.5, indicating that this measurement will have significant distinguishing power between the models.

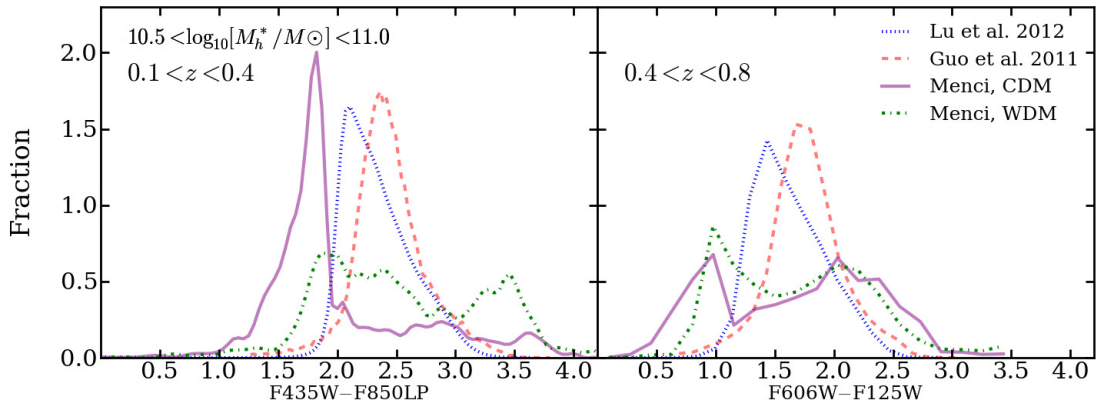

Figure 7.3: Model predictions for the distribution of satellite colors transformed from Figure 4.5 into HST filters selected to straddle the 4000 Å break in each redshift interval. Our measurement of satellite colors in CANDELS imaging will be able to distinguish between these models.

To perform this measurement, we will jointly infer the number, spatial distribution,

and colors of satellites using both the full COSMOS F814 field and the multi band CANDELS fields. The large single-band field constrain the satellite spatial distribution, while the CANDELS fields will provide additional depth and color. We will account for the fact that the spatial distribution of satellites may vary with magnitude in our analysis.

Formally, the probability of measuring $N^{\rm obs}$ objects with positions $\boldsymbol{x}_i$, magnitudes $\boldsymbol{m}_i$ and colors $\{c_i\}$ around the jth host with stellar mass and redshift described by h_j, given a set of model parameters $\boldsymbol{\theta}$ is given by product over the sum of all of the likelihoods of the object having those colors, magnitudes and positions if it were a satellite and if it were a background/foreground object.

$$\begin{aligned}\Pr(\boldsymbol{d}_j|\boldsymbol{\theta},\boldsymbol{h}_j) = \; & \Pr(N^{\rm obs}|\boldsymbol{\theta}) \textstyle\prod_i \Pr(\boldsymbol{x}_i,\boldsymbol{m}_i,\{c_i\}|\boldsymbol{\theta},\boldsymbol{h}_j,S)\Pr(S|\boldsymbol{\theta},\boldsymbol{h}_j)+ \\ & \Pr(\boldsymbol{x}_i,\mathbf{m_i},\{\mathbf{c_i}\}|\boldsymbol{\theta},\mathbf{h_j},\mathbf{B})\Pr(\mathbf{B}|\boldsymbol{\theta},\mathbf{h_j})\end{aligned} \tag{7.1}$$

The term $\Pr(\boldsymbol{x}_i,\boldsymbol{m}_i,\{c_i\}|\boldsymbol{\theta},\boldsymbol{h}_j,S)$ is the probability of finding a satellite at a certain position is described by the the spatial distribution, which we parametrize as a radial power-law, plus elliptical angular distribution while the color distribution of satellites is parametrized as a skewed Normal distribution (N13), and the luminosity function is described by a power-law. The best-fit parameters of these functions will be inferred via a Monte-Carlo sampling of the likelihood function described above.

The term $\Pr(\boldsymbol{x}_i,\boldsymbol{m}_i,\{c_i\}|\boldsymbol{\theta},\boldsymbol{h}_j,B)$ is the probability of finding a background/foreground object at a certain position (which is uniform), with a certain magnitude and colors. The parameters for these three functions will be constrained using measurements of the local

field properties.

The term $\Pr(S|\boldsymbol{\theta}, \boldsymbol{h}_j)$ is the probability of an object being a satellite given only the model parameters,

$$\Pr(S|\boldsymbol{\theta}, \boldsymbol{h}_j) = \frac{N_s}{N_s + N_b} \tag{7.2}$$

and $\Pr(B|\boldsymbol{\theta}, \boldsymbol{h}_j)$ is defined in an analogous way for the line-of-sight interlopers.

The deep IR imaging in CANDELS also makes it possible to constrain the stellar mass function of satellite galaxies, given the tight correlation between IR luminosity and stellar mass (Bell et al. 2003). The stellar masses of satellite galaxies are useful for comparisons with abundance matching, and can help place new constraints on the growth mechanisms of central galaxies, in addition to providing an important new baseline for theoretical models of star formation in satellite galaxies.

Our team is currently undertaking a measurement of the satellite colors and luminosity function, and is funded by the HST archival grant HST-AR-13271. We plan to begin measuring the satellite stellar mass function in the coming year.

In the more distant future, JWST will enable deep, high resolution measurements which will allow a new regime of measurement of satellite galaxies at intermediate and high redshifts.

7.2 Narrow-line lensing going forward

In Chapter 6 we described our pilot analysis of narrow-line lensing in the strong gravitational lens B1422+231. As we discussed, in order to derive a meaningful constraint on the subhalo mass function it is necessary to both increase the sample of narrow-line

lenses, and to use a more sophisticated gravitational lens model which involves realistic populations of subhalos as well as foreground/background structure. In this section we describe ongoing work to increase the sample of narrow-line lenses as well as future plans to use more realistic theoretical models.

7.2.1 Ongoing data analysis

We have obtained OSIRIS IFU data for five targets total. Of these, three (1422, 1413 and 1138) were observed during good weather conditions, and have sufficiently high signal-to-noise ratio to perform the same analysis as was performed for 1422. Two of the targets, 0924 and 0810 were observed during poorer weather conditions.

Figure 7.4 shows a preliminary extraction of the spectrum of image D. The narrow-line emission is detectable although with a low signal to noise ratio. Recently, the OSIRIS grating was upgrading, improving the total instrument throughput by a factor of two. We have applied for time to re-look at 0924 and 0810 as well as to measure narrow-line lensing in three new systems, 0134, 1330 and 2045. In addition to enabling significantly higher signal-to-noise ratio measurements of [NII] emission, the more isolated [SII] and [OI] emission lines should also be detectable with the new grating in similar exposure times.

As we discussed in Chapter 6, thousands of quad quasar lenses are expected to be discovered in ongoing optical surveys which include DES, LSST and PANSTARRS (Oguri & Marshall 2010). While portions of these fields are accessible to Keck, the most southern targets will require observation from an alternate site. One promising alternative is HST

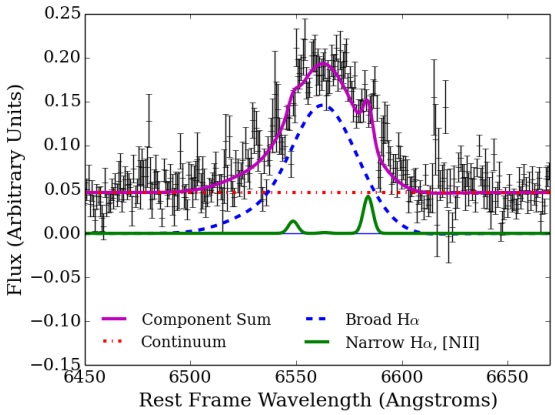

Figure 7.4: The preliminary extracted spectrum of the faintest image of the gravitational lens 0924 with 5400 seconds exposure time in poor conditions. [NII] is detected with a signal to noise ratio of ~ 10. The new grating will make it possible to double the signal-to noise ratio in the [NII] measurement, in addition to significantly detecting narrow [SII] and [OI] emission

WFC3 grism spectroscopy. The space-based imaging provides sufficient spatial resolution to resolve the lensed images, while the spectral resolution is sufficient to resolve Hβ and [OIII] lines. High signal to noise ratio line fluxes can be obtained in about one orbit for most targets brighter than 18th magnitude.

7.3 Combining the results

In this thesis, the subhalo mass function and the properties of luminous satellites have been presented as two separate means of understanding galaxy formation and ΛCDM. An important future step will be combining information from these two probes, which will yield highly complementary information. First, by comparing stellar mass function of satellites around lens-like host galaxies, with the subhalo mass function inferred from the gravitational lensing analysis, it will be possible to test whether there is a significant

population of massive, dark, subhalos. This can be compared with results from Boylan-Kolchin et al. (2012), for instance, who noted that the most luminous Milky Way satellite galaxies exist in subhalos with significantly lower characteristic velocities than predicted by simulations, possibly indicating the presence of massive, dark subhalos. Secondly, the measurement of the subhalo mass function will place important constraints on models of star formation in satellite galaxies by potentially ruling out cosmological models such as warm dark matter which do not contain low mass subhalos (Menci et al. 2012). Finally, some simulations have demonstrated that baryonic effects such as continuous bursts of supernovae can significantly flatten the density profile of dark matter halos of low mass galaxies which would in turn affect their lensing signature (Teyssier et al. 2013; Zolotov et al. 2012; Brooks et al. 2013). By measuring satellite stellar masses and spectral energy distributions at intermediate redshifts, we will be able to test whether these models are feasible which will allow us to test a physically motivated range of subhalo density profiles in our lensing analysis.

Bibliography

Agustsson, I., & Brainerd, T. G. 2010, ApJ, 709, 1321

Akaike, H. 1974, Automatic Control, IEEE Transactions on, 19, 716

Arneson, R. A., Brownstein, J. R., & Bolton, A. S. 2012, ApJ, 753, 4

Aubert, D., Pichon, C., & Colombi, S. 2004, MNRAS, 352, 376

Auger, M. W., Treu, T., Bolton, A. S., et al. 2009, ApJ, 705, 1099

—. 2010, ApJ, 724, 511

Azzalini, A. 1985, Scandinavian Journal of Statistics, 12, 171

Bailin, J., Power, C., Norberg, P., Zaritsky, D., & Gibson, B. K. 2008, MNRAS, 390, 1133

Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ, 600, 681

Battaner, E., & Florido, E. 2000, FCP, 21, 1

Behroozi, P., Wechsler, R., & Wu, H.-Y. 2012, Astrophysics Source Code Library, 10008

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Bell, E. F., Phleps, S., Somerville, R. S., et al. 2006, ApJ, 652, 270

Benítez, N., Ford, H., Bouwens, R., et al. 2004, ApJS, 150, 1

Bennert, N., Falcke, H., Schulz, H., Wilson, A. S., & Wills, B. J. 2002, ApJL, 574, L105

Benson, A. J. 2010, Phys. Rep., 495, 33

Benson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2002, MNRAS, 333, 177

Benson, A. J., Farahi, A., Cole, S., et al. 2013, MNRAS, 428, 1774

Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, ApJ, 697, 1290

Binney, J., & Tremaine, S. 1987, Galactic dynamics

Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, ApJ, 301, 27

Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., & Moustakas, L. A. 2005, ApJ, 624, L21

—. 2006, ApJ, 638, 703

Bolton, A. S., Burles, S., Schlegel, D. J., Eisenstein, D. J., & Brinkmann, J. 2004, AJ, 127, 1860

Bolton, A. S., Treu, T., Koopmans, L. V. E., et al. 2008, ApJ, 684, 248

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440

Bournaud, F., Jog, C. J., & Combes, F. 2007, A&A, 476, 1179

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, ApJL, 752, L5

Boyarsky, A., Lesgourgues, J., Ruchayskiy, O., & Viel, M. 2009, Physical Review Letters, 102, 201304

Boylan-Kolchin, M., Bullock, J. S., & Kaplinghat, M. 2012, MNRAS, 422, 1203

Boylan-Kolchin, M., Ma, C., & Quataert, E. 2008, MNRAS, 383, 93

Boylan-Kolchin, M., & Ma, C.-P. 2007, MNRAS, 374, 1227

Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2006, MNRAS, 369, 1081

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, MNRAS, 398, 1150

Bradač, M., Schneider, P., Steinmetz, M., et al. 2002, A&A, 388, 373

Brainerd, T. G. 2005, ApJL, 628, L101

Brainerd, T. G., Smail, I., & Mould, J. 1995, MNRAS, 275, 781

Brooks, A. M., Kuhlen, M., Zolotov, A., & Hooper, D. 2013, ApJ, 765, 22

Brooks, A. M., & Zolotov, A. 2014, ApJ, 786, 87

Browne, I. W. A., Wilkinson, P. N., Jackson, N. J. F., et al. 2003, MNRAS, 341, 13

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

Budzynski, J. M., Koposov, S., McCarthy, I. G., McGee, S. L., & Belokurov, V. 2012, ArXiv e-prints, arXiv:1201.5491

Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, ApJ, 539, 517

Bundy, K., Ellis, R. S., & Conselice, C. J. 2005, VizieR Online Data Catalog, 7246, 0

Bundy, K., Fukugita, M., Ellis, R. S., et al. 2009, ApJ, 697, 1369

Bundy, K., Treu, T., & Ellis, R. S. 2007, ApJL, 665, L5

Busha, M. T., Wechsler, R. H., Behroozi, P. S., et al. 2011, ApJ, 743, 117

Carroll, S. M. 2004, Spacetime and geometry. An introduction to general relativity

Cassata, P., Guzzo, L., Franceschini, A., et al. 2007, ApJS, 172, 270

Chen, J. 2008, A&A, 484, 347

Chen, J., Kravtsov, A. V., Prada, F., et al. 2006, ApJ, 647, 86

Chiba, M., Minezaki, T., Kashikawa, N., Kataza, H., & Inoue, K. T. 2005, ApJ, 627, 53

Choi, J.-H., Weinberg, M. D., & Katz, N. 2009, MNRAS, 400, 1247

Clowe, D., Gonzalez, A., & Markevitch, M. 2004, ApJ, 604, 596

Colín, P., Avila-Reese, V., & Valenzuela, O. 2000, ApJ, 542, 622

Conroy, C., & Wechsler, R. H. 2009, ApJ, 696, 620

Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2007, ApJ, 668, 826

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

Dalal, N., & Kochanek, C. S. 2002, ApJ, 572, 25

De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2

de Rijcke, S., Penny, S. J., Conselice, C. J., Valcke, S., & Held, E. V. 2009, MNRAS, 393, 798

Deason, A. J., McCarthy, I. G., Font, A. S., et al. 2011, MNRAS, 415, 2607

Dobke, B. M., & King, L. J. 2006, A&A, 460, 647

Dobler, G., & Keeton, C. R. 2006, MNRAS, 365, 1243

Dobler, G., Keeton, C. R., Bolton, A. S., & Burles, S. 2008, ApJ, 685, 57

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

D'Onghia, E., Springel, V., Hernquist, L., & Keres, D. 2010, ApJ, 709, 1138

Dressler, A. 1980, ApJ, 236, 351

Dubinski, J. 1994, ApJ, 431, 617

Dunkley, J., Spergel, D. N., Komatsu, E., et al. 2009, ApJ, 701, 1804

Dutton, A. A., Conroy, C., van den Bosch, F. C., Prada, F., & More, S. 2010, MNRAS, 407, 2

Evans, N. W., & Witt, H. J. 2003, MNRAS, 345, 1351

Fadely, R., & Keeton, C. R. 2012, MNRAS, 419, 936

Fakhouri, O., Ma, C., & Boylan-Kolchin, M. 2010, MNRAS, 406, 2267

Faltenbacher, A., Jing, Y. P., Li, C., et al. 2008, ApJ, 675, 146

Faltenbacher, A., Li, C., Mao, S., et al. 2007, ApJL, 662, L71

Fassnacht, C. D., Koopmans, L. V. E., & Wong, K. C. 2011, MNRAS, 410, 2167

Faure, C., Kneib, J.-P., Covone, G., et al. 2008, ApJS, 176, 19

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Gao, L., White, S. D. M., Jenkins, A., Stoehr, F., & Springel, V. 2004, MNRAS, 355, 819

Garrison-Kimmel, S., Rocha, M., Boylan-Kolchin, M., Bullock, J., & Lally, J. 2013, ArXiv e-prints, arXiv:1301.3137

Gavazzi, R., Treu, T., Rhodes, J. D., et al. 2007, ApJ, 667, 176

Geha, M., Blanton, M. R., Yan, R., & Tinker, J. L. 2012, ApJ, 757, 85

Gehrels, N. 1986, ApJ, 303, 336

Giavalisco, M., et al. 2004, ApJL, 600, L93

Gnedin, O. Y., Hernquist, L., & Ostriker, J. P. 1999, ApJ, 514, 109

Governato, F., Zolotov, A., Pontzen, A., et al. 2012, MNRAS, 422, 1231

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35

Guerras, E., Mediavilla, E., Jimenez-Vicente, J., et al. 2013, ApJ, 764, 160

Guo, Q., Cole, S., Eke, V., & Frenk, C. 2011a, MNRAS, 1278

—. 2012, ArXiv e-prints, arXiv:1201.1296

Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011b, MNRAS, 413, 101

Hayashi, K., & Chiba, M. 2012, ApJ, 755, 145

Hezaveh, Y., Dalal, N., Holder, G., et al. 2013, ApJ, 767, 9

Hopkins, P. F., Bundy, K., Croton, D., et al. 2010, ApJ, 715, 202

Hoyle, B., Masters, K. L., Nichol, R. C., & Bamford, S. P. 2011, ArXiv e-prints, arXiv:1110.6320

Hurvich, C. M., & Tsai, C.-L. 1989, Biometrika, 76, 297

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644

Jackson, N., Bryan, S. E., Mao, S., & Li, C. 2010, MNRAS, 403, 826

Jarosik, N., Bennett, C. L., Dunkley, J., et al. 2011, ApJS, 192, 14

Kamada, A., Yoshida, N., Kohri, K., & Takahashi, T. 2013, ArXiv e-prints, arXiv:1301.2744

Kamionkowski, M., & Liddle, A. R. 2000, Physical Review Letters, 84, 4525

Kang, X., Macciò, A. V., & Dutton, A. A. 2012, ArXiv e-prints, arXiv:1208.0008

Kauffmann, G., White, S. D. M., Heckman, T. M., et al. 2004, MNRAS, 353, 713

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33

Kaufmann, T., Bullock, J. S., Maller, A., & Fang, T. 2008, in American Institute of Physics Conference Series, Vol. 1035, The Evolution of Galaxies Through the Neutral Hydrogen Window, ed. R. Minchin & E. Momjian, 147–150

Kaviraj, S., Peirani, S., Khochfar, S., Silk, J., & Kay, S. 2009, MNRAS, 394, 1713

Kaviraj, S., Tan, K.-M., Ellis, R. S., & Silk, J. 2011, MNRAS, 411, 2148

Kawano, Y., Oguri, M., Matsubara, T., & Ikeuchi, S. 2004, PASJ, 56, 253

Kazantzidis, S., Abadi, M. G., & Navarro, J. F. 2010, ApJL, 720, L62

Keeton, C. R. 2001a, ArXiv Astrophysics e-prints, astro-ph/0102341

—. 2001b, ArXiv Astrophysics e-prints, astro-ph/0102340

—. 2001c, ArXiv Astrophysics e-prints, astro-ph/0111595

Keeton, C. R., Burles, S., Schechter, P. L., & Wambsganss, J. 2006, ApJ, 639, 1

Keeton, C. R., & Moustakas, L. A. 2009, ApJ, 699, 1720

Keeton, C. R., & Zabludoff, A. I. 2004, ApJ, 612, 660

Kelly, B. C. 2007, ApJ, 665, 1489

Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A., & Uttley, P. 2014, ArXiv e-prints, arXiv:1402.5978

Kennicutt, Jr., R. C. 1998, ApJ, 498, 541

Kistler, M. D., Yuksel, H., & Hopkins, A. M. 2013, ArXiv e-prints, arXiv:1305.1630

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82

Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, ApJ, 740, 102

Knebe, A., Gill, S. P. D., Gibson, B. K., et al. 2004, ApJ, 603, 7

Knobel, C., Lilly, S. J., Kovac, K., et al. 2012, ArXiv e-prints, arXiv:1211.5607

Kochanek, C. S. 2002, in The Shapes of Galaxies and their Dark Halos, ed. P. Natarajan, 62–71

Kochanek, C. S., & Dalal, N. 2004, ApJ, 610, 69

Kochanek, C. S., Falco, E. E., Impey, C. D., et al. 2000, ApJ, 543, 131

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, ApJS, 180, 330

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Koopmans, L. V. E. 2005, MNRAS, 363, 1136

Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, 649, 599

Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009a, ApJL, 703, L51

—. 2009b, ApJL, 703, L51

Koposov, S., Belokurov, V., Evans, N. W., et al. 2008, ApJ, 686, 279

Kravtsov, A. 2010, Advances in Astronomy, 2010, arXiv:0906.3295

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004a, ApJ, 609, 35

Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004b, ApJ, 609, 482

Kundic, T., Hogg, D. W., Blandford, R. D., et al. 1997, AJ, 114, 2276

Lagattuta, D. J., Fassnacht, C. D., Auger, M. W., et al. 2010, ApJ, 716, 1579

Lares, M., Lambas, D. G., & Domínguez, M. J. 2011, AJ, 142, 13

Larkin, J., Barczys, M., Krabbe, A., et al. 2006, NAR, 50, 362

Le Fèvre, O., Abraham, R., Lilly, S. J., et al. 2000, MNRAS, 311, 565

Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, ApJ, 744, 159

Libeskind, N. I., Yepes, G., Knebe, A., et al. 2010, MNRAS, 401, 1889

Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70

Liu, L., Gerke, B. F., Wechsler, R. H., Behroozi, P. S., & Busha, M. T. 2011, ApJ, 733, 62

Lu, Y., Mo, H. J., Katz, N., & Weinberg, M. D. 2012, MNRAS, 421, 1779

Lu, Y., Mo, H. J., Weinberg, M. D., & Katz, N. 2011, MNRAS, 416, 1949

Macciò, A. V. 2010, in American Institute of Physics Conference Series, Vol. 1240, American Institute of Physics Conference Series, ed. V. P. Debattista & C. C. Popescu, 355–358

Macciò, A. V., Dutton, A. A., & van den Bosch, F. C. 2008, MNRAS, 391, 1940

Macciò, A. V., Moore, B., Stadel, J., & Diemand, J. 2006, MNRAS, 366, 1529

MacLeod, C. L., Jones, R., Agol, E., & Kochanek, C. S. 2013, ApJ, 773, 35

MacLeod, C. L., Kochanek, C. S., & Agol, E. 2009, ApJ, 699, 1578

Madau, P., Diemand, J., & Kuhlen, M. 2008, ApJ, 679, 1260

Mandelbaum, R., Hirata, C. M., Broderick, T., Seljak, U., & Brinkmann, J. 2006, MNRAS, 370, 1008

Mandelbaum, R., van de Ven, G., & Keeton, C. R. 2009, MNRAS, 398, 635

Mao, S., Jing, Y., Ostriker, J. P., & Weller, J. 2004, ApJL, 604, L5

Mao, S., & Schneider, P. 1998, MNRAS, 295, 587

Max, C., McGrath, E., Gavel, D., et al. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7015, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

McCully, C., Keeton, C. R., Wong, K. C., & Zabludoff, A. I. 2014a, ArXiv e-prints, arXiv:1401.0197

—. 2014b, ArXiv e-prints, arXiv:1401.0197

McKean, J. P., Koopmans, L. V. E., Flack, C. E., et al. 2007, MNRAS, 378, 109

McMillan, P. J. 2011, MNRAS, 414, 2446

Menci, N., Fiore, F., & Lamastra, A. 2012, MNRAS, 421, 2384

Metcalf, R. B., & Amara, A. 2012, MNRAS, 419, 3414

Metcalf, R. B., & Madau, P. 2001, ApJ, 563, 9

Metcalf, R. B., Moustakas, L. A., Bunker, A. J., & Parry, I. R. 2004, ApJ, 607, 43

Metz, M., Kroupa, P., & Jerjen, H. 2009, MNRAS, 394, 2223

Moore, B., Ghigna, S., Governato, F., et al. 1999, ApJL, 524, L19

Morganson, E., & Blandford, R. 2009, MNRAS, 398, 769

Moustakas, L. A., & Metcalf, R. B. 2003, MNRAS, 339, 607

Muñoz, J. A., Falco, E. E., Kochanek, C. S., et al. 1998, Ap&SS, 263, 51

Muñoz, J. A., Kochanek, C. S., & Keeton, C. R. 2001, ApJ, 558, 657

Müller-Sánchez, F., Prieto, M. A., Hicks, E. K. S., et al. 2011, ApJ, 739, 69

Murayama, T., Taniguchi, Y., Evans, A. S., et al. 1999, AJ, 117, 1645

Myers, S. T., Jackson, N. J., Browne, I. W. A., et al. 2003, MNRAS, 341, 1

Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJL, 699, L178

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563

Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162

Nierenberg, A. M., Auger, M. W., Treu, T., Marshall, P. J., & Fassnacht, C. D. 2011, ApJ, 731, 44

Nierenberg, A. M., Auger, M. W., Treu, T., et al. 2012, ApJ, 752, 99

Nipoti, C., Treu, T., Auger, M. W., & Bolton, A. S. 2009, ApJL, 706, L86

Nipoti, C., Treu, T., Leauthaud, A., et al. 2012, ArXiv e-prints, arXiv:1202.0971

Oguri, M., & Marshall, P. J. 2010, MNRAS, 405, 2579

Oke, J. B. 1974, ApJS, 27, 21

Paardekooper, J.-P., Khochfar, S., & Dalla Vecchia, C. 2013, MNRAS, 429, L94

Papastergis, E., Martin, A. M., Giovanelli, R., & Haynes, M. P. 2011, ApJ, 739, 38

Pasquali, A., Gallazzi, A., Fontanot, F., et al. 2010, MNRAS, 407, 937

Patnaik, A. R., Browne, I. W. A., Walsh, D., Chaffee, F. H., & Foltz, C. B. 1992, MNRAS, 259, 1P

Patnaik, A. R., Kemball, A. J., Porcas, R. W., & Garrett, M. A. 1999, MNRAS, 307, L1

Patton, D. R., & Atfield, J. E. 2008, ApJ, 685, 235

Peebles, P. J. E. 1974, A&A, 32, 197

Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, ApJ, 675, 234

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, ArXiv e-prints, arXiv:1303.5076

Polisensky, E., & Ricotti, M. 2011, Phys. Rev. D, 83, 043506

Postman, M., & Geller, M. J. 1984, ApJ, 281, 95

Purcell, C. W., & Zentner, A. R. 2012, JCAP, 12, 7

Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. 2012, ArXiv e-prints, arXiv:1207.2160

Robaina, A. R., Bell, E. F., van der Wel, A., et al. 2010, ApJ, 719, 844

Romano-Díaz, E., Shlosman, I., Heller, C., & Hoffman, Y. 2009, ApJ, 702, 1250

Rusin, D., & Kochanek, C. S. 2005, ApJ, 623, 666

Schneider, A., Smith, R. E., Maccio, A. V., & Moore, B. 2011, ArXiv e-prints, arXiv:1112.0330

Schneider, P., Kochanek, C. S., & Wambsganss, J. 2006, Gravitational Lensing: Strong, Weak and Micro, ed. Schneider, P., Kochanek, C. S., & Wambsganss, J., doi:10.1007/978-3-540-30310-7

Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, ApJS, 172, 38

Sluse, D., Chantry, V., Magain, P., Courbin, F., & Meylan, G. 2012a, A&A, 538, A99

Sluse, D., Hutsemékers, D., Courbin, F., Meylan, G., & Wambsganss, J. 2012b, A&A, 544, A62

Smith, R. E., & Markovic, K. 2011, Phys. Rev. D, 84, 063507

Somerville, R. S. 2002, ApJL, 572, L23

Sonnenfeld, A., Nipoti, C., & Treu, T. 2014, ApJ, 786, 89

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175

Springel, V. 2010, ARAA, 48, 391

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Strigari, L. E. 2013, Physics Reports, 531, 1, galactic searches for dark matter

Strigari, L. E., Bullock, J. S., Kaplinghat, M., et al. 2007, ApJ, 669, 676

Strigari, L. E., & Wechsler, R. H. 2012, ApJ, 749, 75

Sugai, H., Kawai, A., Shimono, A., et al. 2007, ApJ, 660, 1016

Tal, T., Wake, D. A., & van Dokkum, P. G. 2012, ArXiv e-prints, arXiv:1201.5114

Tal, T., Wake, D. A., van Dokkum, P. G., et al. 2011, ArXiv e-prints, arXiv:1108.1392

Teyssier, R., Pontzen, A., Dubois, Y., & Read, J. I. 2013, MNRAS, 493

Tollerud, E. J., Boylan-Kolchin, M., Barton, E. J., Bullock, J. S., & Trinh, C. Q. 2011, ApJ, 738, 102

Tollerud, E. J., Bullock, J. S., Strigari, L. E., & Willman, B. 2008, ApJ, 688, 277

Totsuji, H., & Kihara, T. 1969, PASJ, 21, 221

Treu, T. 2010, ARAA, 48, 87

Treu, T., Ellis, R. S., Kneib, J.-P., et al. 2003, ApJ, 591, 53

Treu, T., Ellis, R. S., Liao, T. X., & van Dokkum, P. G. 2005a, ApJL, 622, L5

Treu, T., Gavazzi, R., Gorecki, A., et al. 2009, ApJ, 690, 670

Treu, T., Ellis, R. S., Liao, T. X., et al. 2005b, ApJ, 633, 174

Trujillo, I., et al. 2006, ApJ, 650, 18

Trujillo-Gomez, S., Klypin, A., Primack, J., & Romanowsky, A. J. 2011, ApJ, 742, 16

Tully, R. B., & Fisher, J. R. 1977, A&A, 54, 661

van Dam, M. A., Bouchez, A. H., Le Mignant, D., et al. 2006, PASP, 118, 310

van de Ven, G., Mandelbaum, R., & Keeton, C. R. 2009, MNRAS, 398, 607

Vegetti, S., Czoske, O., & Koopmans, L. V. E. 2010a, MNRAS, 407, 225

Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2010b, MNRAS, 408, 1969

Vegetti, S., Lagattuta, D. J., McKean, J. P., et al. 2012, Nature, 481, 341

Vegetti et al., S. 2014, MNRAS, submitted

Viel, M., Bolton, J. S., & Haehnelt, M. G. 2009, MNRAS, 399, L39

Villumsen, J. V., Freudling, W., & da Costa, L. N. 1997, ApJ, 481, 578

Wake, D. A., Whitaker, K. E., Labbé, I., et al. 2011, ApJ, 728, 46

Walker, M. G., & Peñarrubia, J. 2011, ApJ, 742, 20

Wang, J., Frenk, C. S., Navarro, J. F., Gao, L., & Sawala, T. 2012, MNRAS, 424, 2715

Wang, W., & White, S. D. M. 2012, MNRAS, 424, 2574

Watson, D. F., Berlind, A. A., McBride, C. K., Hogg, D. W., & Jiang, T. 2012, ApJ, 749, 83

Watson, D. F., Berlind, A. A., McBride, C. K., & Masjedi, M. 2010, ApJ, 709, 115

Weinberg, D. H., Colombi, S., Davé, R., & Katz, N. 2008, ApJ, 678, 6

Williams, R. J., Quadri, R. F., Franx, M., et al. 2010, ApJ, 713, 738

Wizinowich, P. L., Le Mignant, D., Bouchez, A. H., et al. 2006, PASP, 118, 297

Wolf, C., Meisenheimer, K., Kleinheinrich, M., et al. 2004, A&A, 421, 913

Wolf, J., & Bullock, J. S. 2012, ArXiv e-prints, arXiv:1203.4240

Xu, D. D., Mao, S., Cooper, A. P., et al. 2012, MNRAS, 421, 2553

—. 2010, MNRAS, 408, 1721

Xu, D. D., Sluse, D., Gao, L., et al. 2013a, ArXiv e-prints, arXiv:1307.4220

—. 2013b, ArXiv e-prints, arXiv:1307.4220

Xu, D. D., Mao, S., Wang, J., et al. 2009, MNRAS, 398, 1235

Yegorova, I. A., Pizzella, A., & Salucci, P. 2011, A&A, 532, A105

Yoo, J., Kochanek, C. S., Falco, E. E., & McLeod, B. A. 2005, ApJ, 626, 51

—. 2006, ApJ, 642, 22

Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, ApJ, 571, 172

Zentner, A. R. 2006, in EAS Publications Series, Vol. 20, EAS Publications Series, ed. G. A. Mamon, F. Combes, C. Deffayet, & B. Fort, 41–46

Zentner, A. R., & Bullock, J. S. 2003, ApJ, 598, 49

Zentner, A. R., Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2005, ApJ, 629, 219

Zheng, Z., Coil, A. L., & Zehavi, I. 2007, ApJ, 667, 760

Zolotov, A., Brooks, A. M., Willman, B., et al. 2012, ApJ, 761, 71